



# Intentionnalité et interactions motrices : comment appréhender les intentions d'autrui à partir de la dynamique comportementale ?

Daniel Lewkowicz

## ► To cite this version:

Daniel Lewkowicz. Intentionnalité et interactions motrices : comment appréhender les intentions d'autrui à partir de la dynamique comportementale ?. Psychologie. Université Charles de Gaulle - Lille III, 2013. Français. NNT : 2013LIL30038 . tel-01124205

**HAL Id: tel-01124205**

**<https://theses.hal.science/tel-01124205>**

Submitted on 6 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE LILLE NORD DE FRANCE

UNIVERSITE LILLE 3

Unité de Recherche en Sciences Cognitives et Affectives (URECA EA 1059)

# Behavioral and Cognitive basis of Sequential Actions : How to reveal Intentions in Movement Kinematics ?



Thèse en vue de l'obtention du titre de Docteur en Psychologie Cognitive

*par*

Daniel LEWKOWICZ

*sous la direction du*

Pr. Yvonne DELEVOYE-TURELL

*Soutenue le*

6 décembre 2013

## *Composition du Jury :*

Yvonne DELEVOYE-TURELL (Pr.), Université Lille Nord de France (Lille 3) - Directrice

Thierry POZZO (Pr.), Université de Bourgogne - Rapporteur

Philippe GAUSSIER (Pr.), Université Cergy-Pontoise - Rapporteur

Marcel BRASS (Pr.), University of Ghent – Examineur

Yves PAULIGNAN (CR1), Université de Lyon (Lyon 1) - Examineur

Laurent MADELAIN (Pr.), Université Lille Nord de France (Lille 3) – Président



## Table des matières

Remerciements .....	7
Liste des publications : .....	9
Chapitre 1. Introduction Générale.....	11
1.1. Préambule .....	11
1.2. Un substrat neuronal unique pour l'action, l'observation et la simulation motrice : le système miroir .....	23
1.3. Limites et critiques des Neurones Miroirs.....	32
Chapitre 2. Modèles computationnels de l'action et approches méthodologiques .....	36
2.1. Les modèles internes .....	36
2.2. Methodologie expérimentale .....	43
Chapitre 3. Exécution.....	47
3.1. Coupled Planning of sequential Pick and Place under Speed and Accuracy constraints.....	48
Introduction .....	49
EXPERIMENT 1: Real time to monitor compliancy to speed constraints .....	57
Method .....	57
Results and Discussion.....	61
EXPERIMENT 2: Kinematic analysis of the pick-and-place elements.....	66
Method .....	66
Results and Discussion.....	67
EXPERIMENT 3: Temporal predictability required for coupled planning .....	75
Method .....	76
Results and Discussion.....	78

GENERAL DISCUSSION.....	83
3.2. Effet de la variabilité terminale sur la vitesse d'exécution et le couplage des mouvements inclus dans une même séquence motrice. ....	93
Introduction .....	94
Méthodologie .....	101
Résultats .....	104
Chapitre 4. Observation et Modélisation .....	115
4.1. Reading Motor Intention through Mental Imagery. ....	116
Abstract .....	116
Introduction .....	116
Methods .....	121
Part A: Creating The Stimuli.....	121
Part B: Human Prediction Of Ongoing Actions.....	126
Part C: Predicting Sequential Actions Using Artificial Neural Networks .....	130
Discussion .....	132
Conclusion.....	138
4.2. Eye movement strategies to Read Intentions.....	140
Introduction .....	140
Methods.....	142
Results .....	148
Conclusion.....	156
Chapitre 5. Interaction Sociale.....	159
5.1. Effects of social intention on movement kinematics in cooperative actions.....	160

Introduction .....	161
Materials and methods .....	165
Results .....	173
Discussion .....	179
Chapitre 6. Discussion Générale .....	186
6.1. Rappel des principaux résultats .....	186
6.2. Conséquences pour les interactions sociales .....	196
6.3. Perspectives en Neurosciences .....	198
6.4. Perspectives en Psychologie Cognitive .....	201
6.5. Perspectives en modèles computationnels pour la Robotique .....	203
Références .....	207



# Remerciements

Je tiens tout d'abord à remercier ma directrice de thèse, Pr. Yvonne DELEVOYE-TURELL qui m'a fait confiance tout au long de ces trois années, et qui a su me motiver en me faisant travailler sur mes faiblesses tout en me permettant de réaliser des choses passionnantes et en mettant à ma disposition à la fois un équipement de pointe et du temps pour m'écouter, m'aider, me relire, me conseiller et m'encourager. Yvonne est le genre de personne qu'on ne rencontre que rarement dans une vie, et on en revient toujours grandit.

Mes remerciements vont également à messieurs les professeurs Philippe GAUSSIER et Thierry POZZO, qui m'ont fait l'honneur d'accepter d'être rapporteurs de cette thèse ainsi qu'à Yves PAULIGNAN, et Marcel BRASS en tant qu'examineur et Laurent MADELAIN en tant que président du jury. Je tiens à remercier particulièrement Philippe pour avoir été le coordinateur du projet INTERACT et pour les discussions d'une rare richesse que nous avons pu avoir avec tous les collaborateurs du projet.

Je remercie aussi mes collègues et amis du laboratoire URECA, avec qui j'ai passé d'excellents moments de partage, de discussions plus ou moins sérieuses autour d'une bonne bière des Flandres ou de jeux de plateau endiablés. Il m'est difficile de tous les citer, mais mes pensées vont à Antoine, Armelle, Arthur, Aurelien, Caroline, Célia, Christelle, Delphine, Emanuelle, Emilie(s), Eve, Ewa, Francis, François, Guillaume, Jérémy, Karinne, Laurent, Lucie, Marie-Charlotte, Marie-Laure, Mariama, Mike, Noëlia, Xavier, Yannick et tous ceux que j'oublie.

Je tiens aussi à remercier les membres du laboratoire URECA, Angela, Angèle, Clémence, Fabien, Jean-Louis, Jeremie, Laurent(s), Laetitia, Séverine, Yann et les autres pour leur efforts afin de créer un environnement de travail agréable et efficace pour les doctorants.



Ce travail n'aurait toutefois jamais vu le jour sans la contribution des nombreux étudiants stagiaires que j'ai eu la chance d'encadrer ou avec qui j'ai pu collaborer. En particulier Manon, David, Anthony, Sébastien, Nicolas et Kevin.

Mes pensées vont également à ma famille, Marie-Hélène, Jacques, Myriam, David, Emmanuel et mes trois petits neveux. Merci de m'avoir soutenu et de m'avoir transmis la passion pour la recherche scientifique.

Enfin je remercie ma tendre Anne-Lise qui a réussi à me supporter pendant les moments difficiles et avec qui je partage tant de joie et de bonheur. Tu as fait le chemin de Toulouse à Lille avec moi et c'est avec toute ma reconnaissance et mon admiration que je te dédicace cette thèse.

# Liste des publications :

## Articles publiés :

Lewkowicz, D., Delevoye-Turrell, Y. N., Bailly, D., Andry, P., & Gaussier, P. (2013). Reading Motor Intention through Mental Imagery. *Adaptive Behavior : Special Issue on Artificial Mental Imagery in Cognitive Systems and Robotics*.

Quesque, F., Lewkowicz, D., Delevoye-Turrell, Y. N., & Coello, Y. (2013). Effects of social intention on movement kinematics in cooperative actions. *Frontiers in Neurorobotics*, 7, 14.

## Articles en préparation :

Lewkowicz, D., Delevoye-Turrell, Y. N. (submitted) Coupled Planning under Speed and Accuracy Constraints. *Journal of Experimental Psychology : General*

Lewkowicz, D., Delevoye-Turrell, Y. N. (in prep.) Eye Movement Strategy for Reading Intentions. *Journal of Experimental Psychology : Human Perception and Performance*

## Actes de conférences :

Lewkowicz, D., Delevoye-Turrell, Y. (2013) Reading Motor Intentions. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35<sup>th</sup> Annual Conference of the Cognitive Science Society* (pp. 2872-2877). Austin, TX: Cognitive Science Society.

Lewkowicz, D., Delevoye-Turrell, Y. (2012) Effet de la variabilité terminale sur la vitesse d'exécution et le couplage de mouvements séquentiels. *Mosaïque n°10 – ACTES : « 7<sup>ème</sup> Journée Scientifique des Jeunes Chercheurs en Psychologie »* (p. 35-55). Lille, France.

**Communications orales :**

Lewkowicz, D., Delevoye-Turrell, Y. (Septembre 2012) Planification couplée d'une tâche séquentielle sous contrainte de vitesse et de précision. Communication orale présentée au 54ème congrès de la Société Française de Psychologie.

Lewkowicz, D., Delevoye-Turrell, Y. (Juin 2012) Effet de la variabilité terminale sur la vitesse d'exécution et le couplage de mouvements séquentiels. Communication orale présentée à la 7ème Journée des Jeunes Chercheurs en Psychologie.

Lewkowicz, D., Delevoye-Turrell, Y. (Juin 2011) Effet du Contexte sur la Planification de l'Action. Communication orale présentée à la 6ème Journée des Jeunes Chercheurs en Psychologie.

**Communications affichées :**

Lewkowicz, D., Delevoye-Turrell, Y. (2013, August) Reading Motor Intentions. Poster presented at the 35th Annual Conference of the Cognitive Science Society. Berlin, Germany.

Lewkowicz, D., Delevoye-Turrell, Y. (Mars 2012) Mouvements humains : Lecture de l'intentionnalité. Les différentes approches en Psychologie, en Sciences Cognitives et en Robotique. Poster présenté au 11ème Forum des Sciences Cognitives.

# Chapitre 1. Introduction Générale

## 1.1. Préambule

### 1.1.1. La robotique d'assistance : un challenge pour le 21ème siècle ?

D'après les propos de Bruno Bonell, Pdg de l'entreprise Robopolis à Lyon, le 21ème siècle verra naître une nouvelle révolution industrielle, avec l'avènement du robot. Ce qu'il appelle la « robolution » serait d'ores et déjà en marche, en particulier dans le rapport des hommes à la machine : « Nous sommes passé d'un phénomène gadget à un phénomène de société (...) Nous ne sommes plus dans un rapport à l'outil, mais dans une relation affective avec quelque chose qui nous ressemble ». (La Tribune, 7 Mai 2013)

Depuis ces 20 dernières années, des évolutions particulièrement importantes ont été réalisées dans le domaine de la robotique, et certains robots sont déjà largement entrés dans notre quotidien (robots nettoyeurs de piscine, aspirateurs domestiques, etc...). Un des domaines de la robotique où les innovations sont parmi les plus développées en France et dans le monde est celui de la Robotique d'Assistance. Et pour cause, selon le département des Affaires Economiques et Sociales des Nations Unies, la population mondiale de personnes âgées de plus de 85 ans va tripler d'ici 2050, soit plus de 180 millions de personnes (Estimations « medium fertility » de « World Population Prospects : The 2012 Revision »). Au Japon, où la population de plus de 65 ans a déjà dépassé la barre des 20% depuis 2006, et dont les prévisions amènent ce chiffre à 35% en 2050, il n'est pas étonnant que ce pays soit parmi les plus mobilisés sur la recherche scientifique dans ce domaine.

L'objectif de la robotique d'assistance est de proposer un support physique, psychique ou social pour améliorer l'autonomie des personnes dépendantes. Ce domaine est particulièrement important dans le cadre des interactions homme-robot car il met l'accent sur des interactions de proximité avec des personnes potentiellement en situation de handicap ou d'invalidité. Le challenge à relever pour ce domaine est donc de proposer des contacts

physiques sécurisés mais aussi un véritable aspect social, à travers une architecture cognitive et émotionnelle et des interactions naturelles comme des gestes ou des paroles (Goodrich & Schultz, 2007). Ainsi, au-delà des capacités « classiques » telles que bouger ou agir de manière autonome, la robotique d'assistance se concentre sur comment utiliser au mieux l'aspect physique et incarné du robot pour communiquer ou interagir de manière intéressante pour autrui (Tapus, Mataric, & Scassellati, 2007). L'étude des interactions homme-robot pour la robotique d'assistance est un champ inter-disciplinaire par essence qui met en relation de nombreux domaines tels que la robotique, la médecine, les sciences sociales et cognitives, les neurosciences ou la psychologie développementale.

Une distinction importante reste néanmoins à faire entre la robotique d'assistance dite « de contact » qui se concentre sur les aspects physiques des interactions de proximité comme proposé ci-dessus et la robotique d'assistance « sociale » (SAR) qui cherche à aider l'utilisateur aussi bien que de le divertir et l'accompagner au travers des interactions sociales. De plus, le terme « socialement interactif » a été introduit pour distinguer les interactions sociales réelles de la simple télé-opération dans le cadre plus général des interactions hommes-robots (HRI) et permet de décrire des robots dont la tâche principale est justement d'établir une interaction (Fong, Nourbakhsh, & Dautenhahn, 2003). La robotique d'assistance sociale (SAR) telle que définie par Feil-Seifer et Mataric en 2005 est en fait à l'intersection des champs disciplinaires que sont la robotique d'assistance (AR) et la robotique socialement interactive (SIR) proposée par Fong (2003).

La robotique d'assistance *sociale* partage donc avec la robotique d'assistance « générale » l'objectif de proposer une assistance aux utilisateurs humains. Mais de surcroît, cette assistance doit se faire à travers une interaction sociale incarnée et en temps réel tel que proposé dans la robotique socialement interactive (SIR). En d'autres termes, dans la SIR l'objectif du robot est de développer des interactions les plus efficaces possibles avec l'humain dans le but *d'établir des interactions*. Alors que dans la SAR l'objectif du robot est

de créer des interactions efficaces dans le but de *porter assistance* et de réaliser des progrès dans la convalescence, la réhabilitation, l'apprentissage, etc... (Feil-Seifer & Mataric, 2005).

Ainsi, Feil-Seifer et Mataric proposent un grand nombre d'applications possibles de la SAR et concluent sur trois critères principaux pour l'évaluation de telles architectures :

1) Les performances de l'utilisateur : un système SAR doit intéresser l'utilisateur pour lui permettre d'atteindre les objectifs proposés tout en répondant à la fois aux objectifs du patient et à ceux des personnels soignants ou de l'équipe d'accompagnement. Un robot peut agir de manière intéressante et attirante pour un utilisateur mais s'avérer complètement inefficace au regard des progrès de cette personne. Une telle architecture doit donc être capable d'être « au service de plusieurs maîtres » dont les objectifs peuvent potentiellement être en conflit ou varier significativement. Ce qui représente déjà un challenge complexe pour une telle architecture.

2) Le niveau d'autonomie : Un système SAR ne doit pas requérir un opérateur expert ou un entraînement intensif pour être utilisé. Il doit être capable de démarrer, s'arrêter et être configuré avec un effort minimum tout en se conformant aux routines changeantes et aux exigences variables des utilisateurs et du personnel.

3) Incarné contre Non incarné : Enfin, un système SAR pose la question du rôle d'une architecture incarnée pour de telles tâches. Si le robot n'a pas besoin de réaliser des tâches incluant des contacts physiques avec l'utilisateur (voir définitions plus haut), pourquoi alors utiliser un véritable robot dans ce cas ? Pourquoi est-ce qu'un ordinateur ou un assistant personnel digitalisé ne pourrait pas être suffisant ? La robotique sociale repose sur l'idée que l'humain a une large tendance à attribuer des buts ou des intentions à de simples entités mobiles. Un des enjeux majeurs sera alors de comprendre comment ces caractéristiques inhérentes à l'homme seraient incluses dans des architectures incarnées pour ensuite être traduites en progrès significatifs en matière d'assistance robotique en thérapie, convalescence ou apprentissage.

Au cours des travaux proposés dans cette thèse, c'est ce dernier point qui a le plus retenu notre attention car il propose des questionnements scientifiques pertinents en robotique mais aussi et surtout en psychologie expérimentale et en sciences cognitives, notamment pour ce qui est du lien entre les caractéristiques externes des mouvements humains et la question de l'attribution des buts et des intentions. Mais avant de rentrer en détail dans les travaux expérimentaux que nous avons mené autour de ce sujet chez l'homme, nous souhaitons proposer au lecteur quelques éclaircissements nécessaires pour bien comprendre les enjeux et la portée de nos résultats. Ainsi dans les paragraphes suivants, une problématique plus précise est proposée à partir de certaines constatations expérimentales puis un cadre théorique commun est proposé pour faire un bref récapitulatif des connaissances actuelles chez l'homme et le robot.

### **1.1.2. L'importance de l'aspect biologique des mouvements**

Au cours des dernières décennies, un très grand nombre de travaux expérimentaux en psychologie expérimentale ont démontré que l'homme était particulièrement sensible aux propriétés biologiques du mouvement. Par exemple, depuis les travaux de Johansson (1973), les chercheurs ont pu démontrer à de multiples reprises que la simple présentation sous forme de points lumineux sur fond noir des positions articulaires d'une personne en train de marcher, pouvait entraîner chez l'observateur la sensation irrésistible et vivace d'une véritable personne en action alors que ce percept s'effondre en une simple image dénuée de toute signification dès lors que le marcheur reste immobile (G. Johansson, 1973). Ces travaux ont permis l'utilisation de stimuli facilement manipulables en terme de forme, de taille ou de propriétés cinématiques pour étudier le phénomène d'attribution d'intention. Par exemple, des batteries d'animation de simples formes géométriques continuent à être largement utilisées en neuro-imagerie chez les personnes atteintes de trouble autistiques (Castelli, Frith, Happé, & Frith, 2002; Kana, Keller, Cherkassky, Minshew, & Just, 2009) et en schizophrénie (Horan et al., 2009; Koelkebeck et al., 2010) pour tester si les patients interprètent ces mouvements de la même manière que des personnes non atteintes. Différents scénarios sont proposés tels que

la poursuite, l'évasion, l'affrontement, la danse ou la séduction et ont été catégorisés selon leurs propriétés cinématiques (Blythe, Todd, & Miller, 1999). Ce qui est frappant, lorsqu'on observe ces animations c'est la capacité automatique et particulièrement rapide de la formation de tels percepts (G. Johansson, 1976) qui suggère que l'homme posséderait des régions cérébrales dont la fonction serait dédiée uniquement à ce type de tâche.

Ainsi, chez l'homme, il a été montré que certaines aires cérébrales spécifiques étaient actives lors de la perception de mouvements humains en particulier dans la région de la jonction pariéto-occipito-temporale (Giese & Poggio, 2003; Grezes et al., 2001; Grossman et al., 2000). Ce qui est particulièrement frappant c'est la sélectivité de ces aires cérébrales pour l'aspect biologique des mouvements. Car en effet, ces mêmes régions ne sont pas actives si les points effectuent des mouvements aléatoires, ou s'ils effectuent les mêmes mouvements mais en partant de positions initiales mélangées ou inversées (Sumi, 1984). Ce phénomène est généralement interprété comme étant le résultat d'un processus de sélection naturelle concernant les capacités de discrimination des mouvements pour des raisons de survie ou de reproduction. Par exemple, une mauvaise lecture d'intention pourrait conduire un animal à interpréter les mouvements d'un prédateur comme ceux d'une parade amoureuse ou ceux d'un jeu bénin comme une rivalité agressive. Par ailleurs, plus ces jugements sont fait rapidement, plus le temps disponible pour effectuer une réponse appropriée sera grand, que ce soit pour fuir, approcher ou suivre l'individu observé. Enfin, le fait que les humains réalisent cette tâche de façon automatique (Scholl & Tremoulet, 2000) suggère que la sélection naturelle aurait favorisé des lectures d'intentions à la fois rapides et précises. Ainsi, il semble alors important de s'intéresser aux détails de la cinématique des mouvements qui sont observés car ils seraient porteurs d'informations.

Plus récemment, et grâce aux progrès récents en robotique humanoïde, des auteurs (Oztop, Franklin, Chaminade, & Cheng, 2005) ont pu comparer directement la perception de mouvements réalisé par un participant humain par rapport à des mouvements effectués par un robot taille humaine. Au cours de cette expérience, les auteurs ont utilisé un paradigme



expérimental simple et désormais largement réutilisé pour répondre à la question suivante : « est-ce que les mouvements d'un robot humanoïde sont perçus comme ceux d'un humain ? » (Kilner, Paulignan, & Blakemore, 2003; Oztop, Franklin, Chaminade, & Cheng, 2005). Leur protocole expérimental propose aux participants de réaliser des mouvements rythmiques du bras tout en observant ceux d'un partenaire humain ou robotique. Les mouvements sont réalisés de manière verticale (de haut en bas) ou horizontale (de gauche à droite). Toute l'astuce du protocole réside dans l'idée de montrer aux participants des mouvements qui sont congruent avec les leurs (même axe) ou incongruents (axe différent), comme présenté dans la Figure 1. Leurs résultats montrent que la variabilité des mouvements est fortement impactée uniquement dans la condition où les sujets observent des humains agissant de manière incongruente (Figure 1). Cet effet d'interférence motrice entre l'action observée et l'action exécutée démontre bien que lorsque les mouvements d'un robot n'entraînent pas les mêmes conséquences que les mouvements biologiques dans le cerveau des observateurs. Les auteurs proposent ici une méthode comportementale pour vérifier l'hypothèse selon laquelle le cerveau traite les informations des mouvements biologiques de manière différente des autres mouvements dans l'environnement.

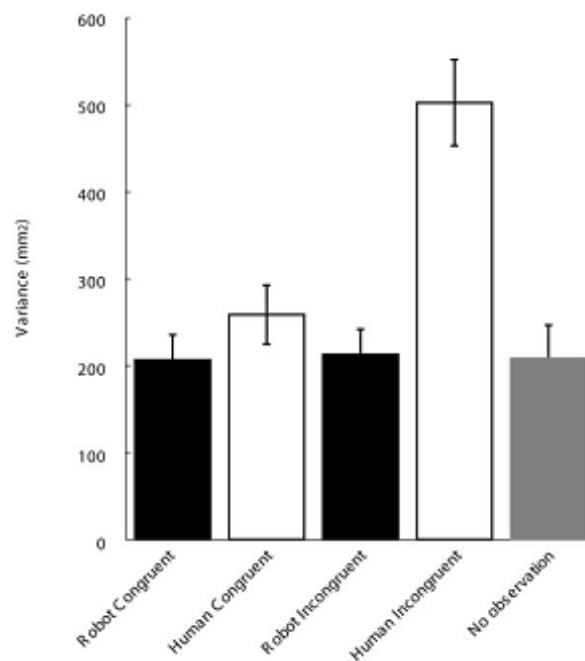
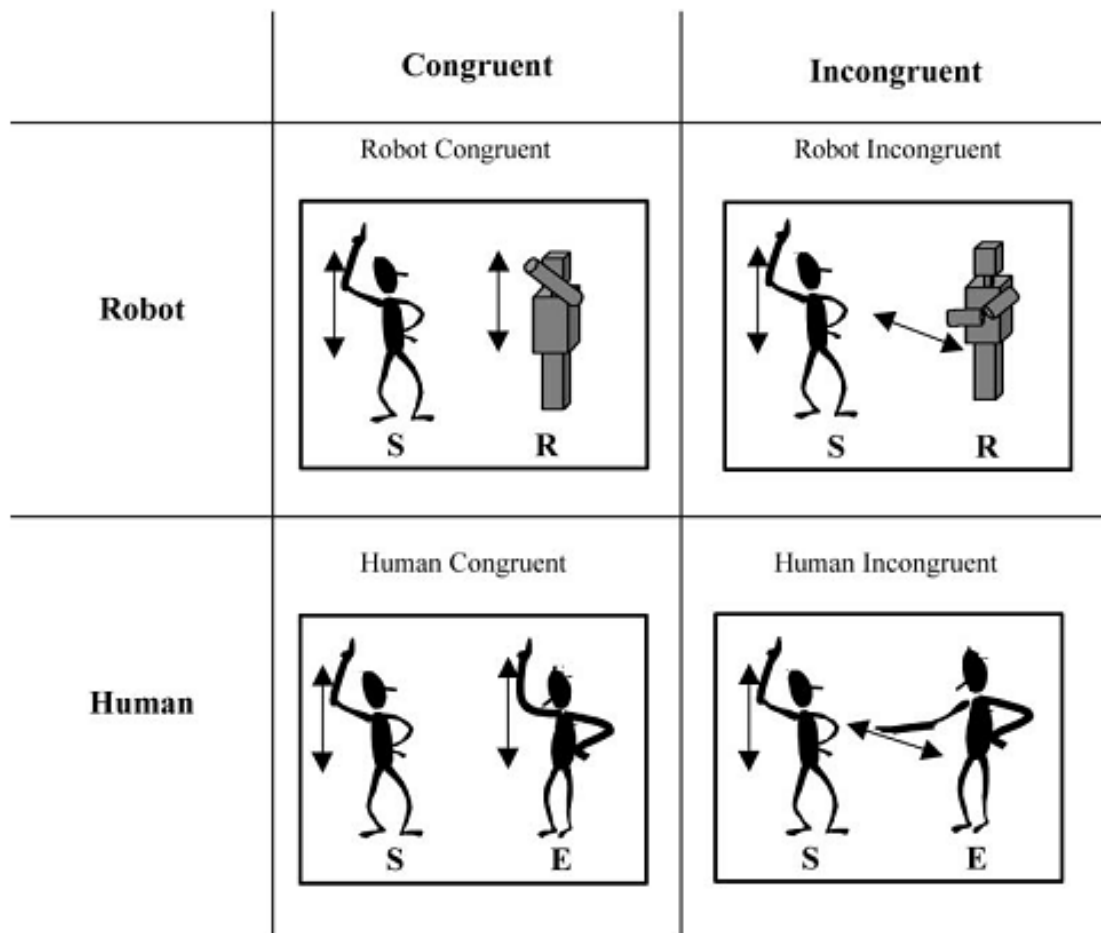


Figure 1. adaptée de Kilner, J.M, Y Paulignan, et S.J Blakemore. «An interference effect of observed biological movement on action. ». *Current Biology* (2003)

Une autre manière d'aborder ce phénomène est de demander aux observateurs d'estimer la position finale d'un stimulus en mouvement et de comparer leur prédiction selon

l'aspect biologique ou non des mouvements du stimulus. C'est l'objet des travaux effectués par Pozzo et collaborateurs en 2006 qui ont montré que les composantes cinématiques du mouvement vont entraîner dans le cas des mouvements non biologiques des erreurs d'estimations plus importantes que pour des mouvements biologiques (Pozzo, Papaxanthis, Petit, Schweighofer, & Stucchi, 2006). De plus, les auteurs reportent une incapacité pour les observateurs à reconnaître de façon explicite si les mouvements possédaient des caractéristiques biologiques ou non, ce qui renforce l'idée d'un processus rapide et automatique pour le traitement de ce type d'information.

Enfin, une autre propriété émergente de ce phénomène est la capacité des observateurs à juger si un mouvement est possible en fonction de ses contraintes spatiales et temporelles. En utilisant un paradigme de mouvement apparent à partir de deux images statiques d'un bras à deux positions différentes, Grosjean et collaborateurs (2007) ont demandé à leurs participants de juger de la possibilité d'effectuer de tels mouvements compte tenu de la précision terminale et de la durée s'écoulant entre la présentation des deux images. Ici encore, l'intérêt de cette expérimentation était de pouvoir vérifier si les mouvements d'un bras robotique étaient jugés selon les mêmes critères que les mouvements d'un bras humain (Figure 2). Leurs résultats montrent qu'effectivement, les humains donnent leur réponse en fonction de leurs connaissances inhérentes des lois biomécaniques du mouvement du bras telle que le trade-off entre vitesse et précision (Grosjean, Shiffrar, & Knoblich, 2007), ou plus communément appelé « Loi de Fitts » en psychologie expérimentale (Fitts, 1954). Ce qui est particulièrement intéressant est le fait que pour les mouvements du robot, les humains semblent utiliser les mêmes lois que pour des humains et donc jugent le comportement du robot en attribuant des caractéristiques comportementales humaines à une entité non-humaine, autrement dit, à faire de l'anthropomorphisme.

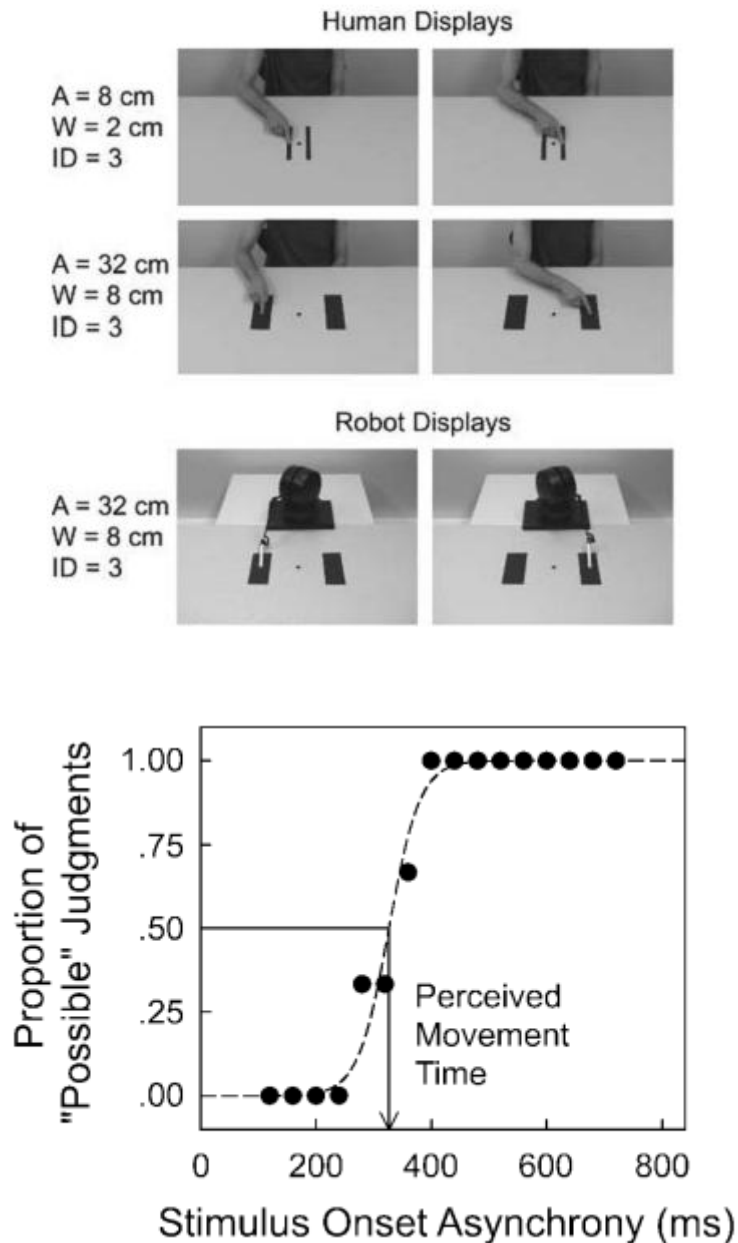


Figure 2. Adaptée de Grosjean, Marc, Maggie Shiffrar, et Günther Knoblich. « Fitts's Law Holds for Action Perception ». *Psychological Science* (2007)

Le phénomène d'anthropomorphisme est d'autant plus pertinent en robotique humanoïde qu'il a donné lieu à certaines constatations intrigantes. Par exemple, un des phénomènes les plus connus en robotique humanoïde est le phénomène de la « vallée dérangement » (*Uncanny Valley*) proposé par Masahiro Mori en 1970 (Figure 3). Ce phénomène peut se résumer ainsi : Plus l'apparence d'un agent est proche de celle d'un humain, plus la disposition des observateurs sera positive à son égard, jusqu'à un point où une

poursuite supplémentaire vers la ressemblance avec l'humain sera considéré comme étrange, non-familière et dérangeante. (Mori, 1970).

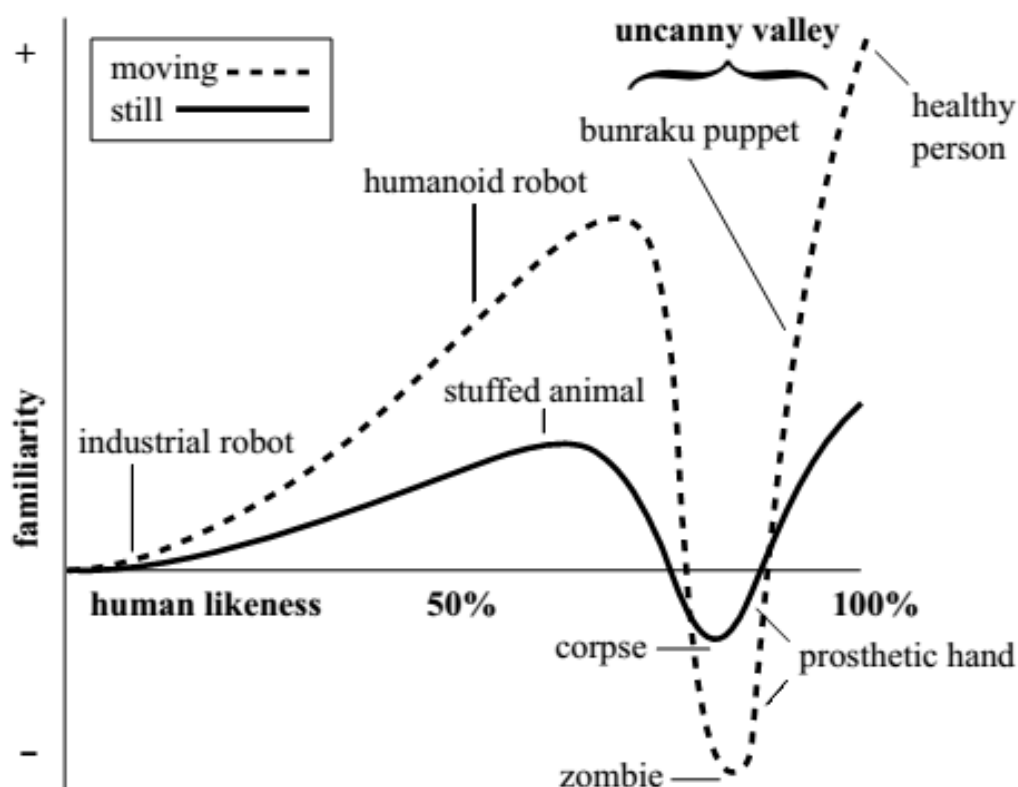


Figure 3. Adaptée de Mori, M.. « The uncanny valley ». *Energy* (1970)

Cette idée a été particulièrement mise en avant ces dernières années, en raison de la popularité des robots réalisés par Hiroshi Ishiguro ou David Hanson dont la qualité de l'aspect externe des visages est particulièrement fidèle aux visages humains (Hanson et al., 2005). Sans entrer dans une polémique qui consisterait à chercher à vérifier si le phénomène de la *uncanny valley* possède une validité sur le plan scientifique, l'observation des mouvements des robots de Hiroshi Isiguro a par exemple permis de faire des constatations intéressantes concernant les processus cérébraux sous-jacents (Saygin, Chaminade, Ishiguro, Driver, & Frith, 2012).

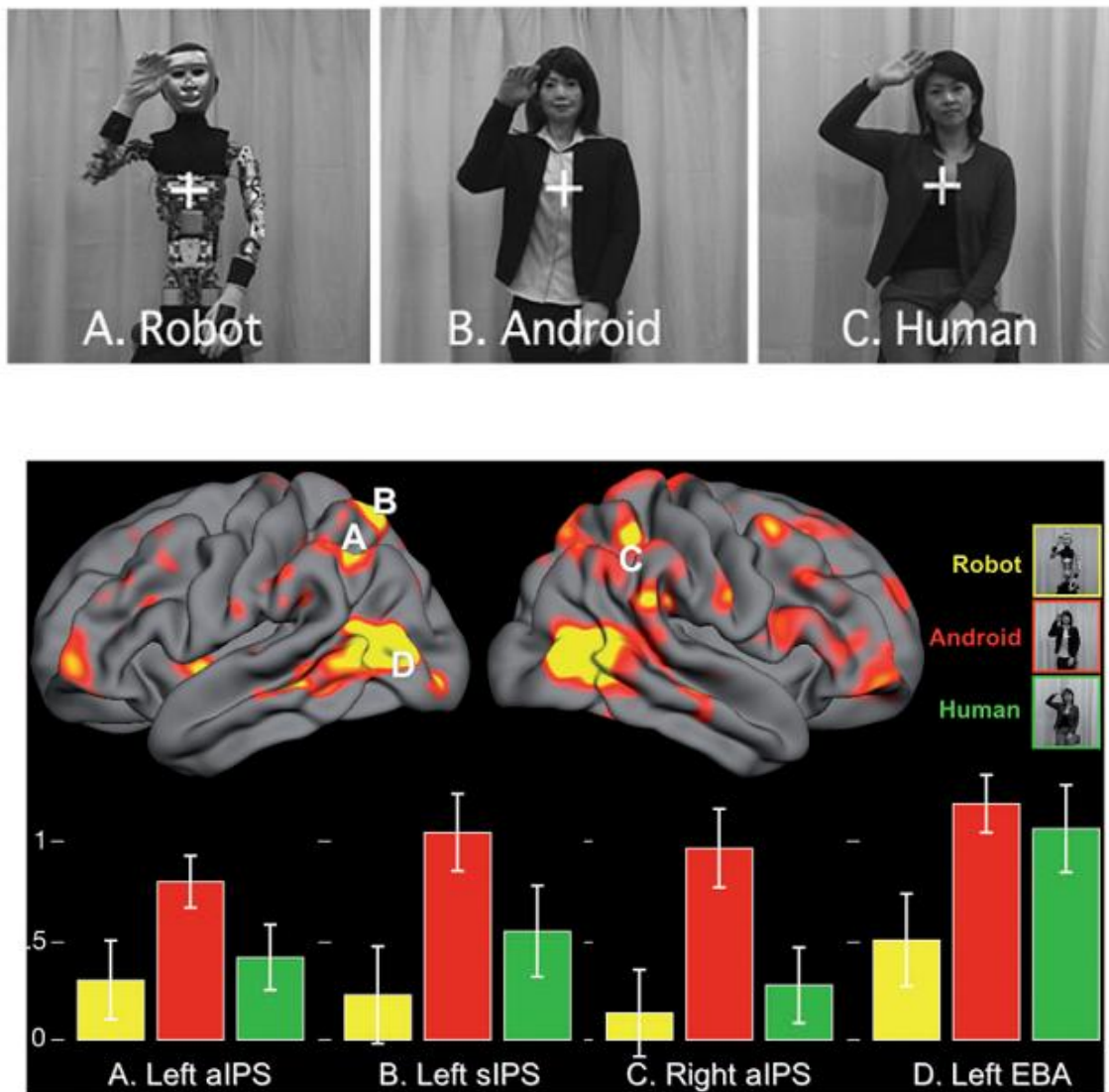


Figure 4. Adaptée de Saygin, Ayse Pinar, Thierry Chaminade, Hiroshi Ishiguro, Jon Driver, et Chris Frith. « The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions ». *Social cognitive and affective neuroscience* (2012)

En effet, en comparant des mouvements effectués par un robot humanoïde à l'aspect robotique avec ceux d'un androïde à l'aspect humain ou ceux d'un véritable humain, Saygin et collaborateurs (2012) ont pu mettre en évidence le fait que la disparité entre l'apparence extérieure et l'aspect biologique des mouvements entraînait dans le cas de l'androïde des activations plus intenses notamment dans les aires pariétales (voir Figure 4). Ces aires sont notamment connues pour leurs propriétés d'intégration sensori-motrice et la présence de « neurones miroirs » sur lesquels nous reviendrons plus en détails dans la section suivante de cette introduction.

Il a alors été proposé que l'effet de la *uncanny valley* était due à une erreur de prédiction motrice. Lorsqu'un agent ressemble fortement à un humain, basé sur l'expérience d'une vie entière, le cerveau génère une prédiction forte que cette apparence va engendrer des comportements spécifiques et des propriétés particulières au niveau cinématique. Le conflit entre les prédictions du cerveau et ce qui est vu en réalité (l'erreur de prédiction) aura comme première conséquence de rendre les mouvements de l'agent non familiers et étranges.

En conséquence, pour être attractif pour l'humain, un robot ne doit pas simplement reproduire l'apparence humaine mais il doit surtout réaliser des mouvements qui ont un aspect biologique et dont les prédictions peuvent être réalisées facilement par l'observateur. Ceci est la formulation explicite de la problématique de cette thèse. Dans le cas particulier des actions séquentielles, est-ce que les caractéristiques biologiques d'un mouvement séquentiel possèdent des similitudes avec les lois observées sur des mouvements unitaires ? Est-il possible qu'un système artificiel puisse discriminer des intentions humaines à travers la simple observation de la cinématique des mouvements ? Comment un robot peut à travers ses mouvements permettre à un observateur de lire ses intentions ? Dans le cas d'une interaction sociale entre deux personnes, est-ce qu'il existe d'autres phénomènes spécifiques aux tâches conjointes qui se reflètent aussi dans la cinématique des mouvements mais qui seraient indépendants de l'intention d'interagir ?

Nous allons à présent nous arrêter quelques instants sur deux notions fondamentales qui ont été abordées brièvement dans les paragraphes précédents mais dont la portée scientifique a permis des avancées majeures dans la compréhension de ces phénomènes en sciences cognitives et en neurosciences computationnelles. Ces deux notions sont 1) l'existence de réseaux de neurones qui partagent des fonctions d'observation et d'exécution, plus connues sous le nom de « neurones miroirs » et 2) les capacités prédictives du système moteur humain et leur modélisation.

## **1.2. Un substrat neuronal unique pour l'action, l'observation et la simulation motrice : le système miroir**

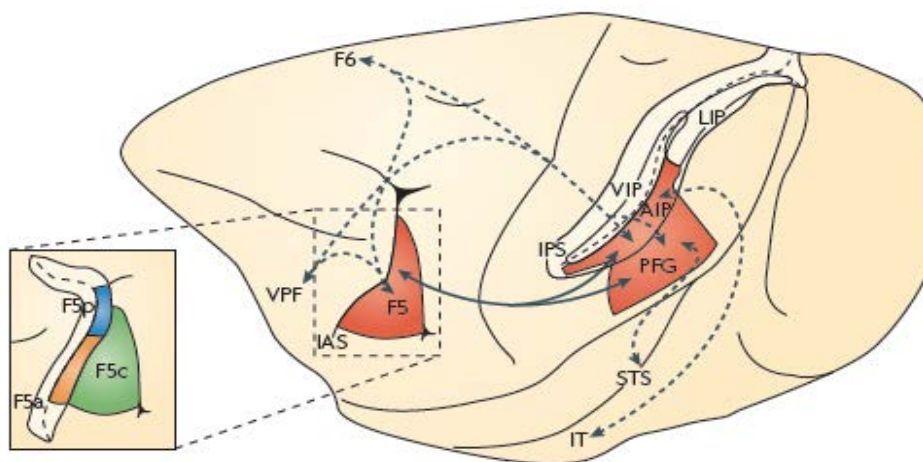
Nous allons dans les paragraphes suivants présenter une des découvertes majeures de ces dernières décennies dans le champ des neurosciences : le système de neurones miroirs, découvert préalablement chez le singe et qui a donné lieu à des propositions concernant des réseaux neuronaux qui partagent des fonctions d'observation et d'exécution chez l'homme.

### **1.2.1. Les études originales chez le singe et les premières données : action et observation**

Les neurones miroirs ont été découverts initialement dans le cortex prémoteur ventral (aire F5) chez le macaque (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Les enregistrements unicellulaires en neurophysiologie ont montré que ces neurones étaient actifs lorsque le singe exécute des actes moteurs mais aussi lorsqu'il observe un autre individu réalisant le même acte. Par la suite, des neurones similaires ont été découverts dans l'aire PFG (Fogassi et al., 2005; Gallese, Fadiga, Fogassi, & Rizzolatti, 2002; Rozzi, Ferrari, Bonini, Rizzolatti, & Fogassi, 2008) et dans l'aire intrapariétale antérieure AIP (Belmalih et al., 2007; Rizzolatti, Fogassi, & Gallese, 2009). Ces deux aires sont largement connectées avec l'aire F5, PFG principalement avec F5c et AIP (Intra-pariétale antérieure) principalement avec F5a (Rozzi et al., 2006). Les deux aires PFG et AIP reçoivent des informations visuelles de haut niveau du cortex situé dans la scissure temporale supérieure (STS) (Rizzolatti et al., 2009; Rozzi et al., 2006). L'aire AIP reçoit quant à elle des projections en provenance du gyrus temporal moyen (Borra et al., 2008). Cette afférence pourrait ainsi fournir aux aires miroirs des informations sur l'identité des objets. Enfin, l'aire F5 est connectée avec l'aire F6, l'aire pré-supplémentaire motrice (pre-SMA) et avec le cortex préfrontal, l'aire 46. Le cortex préfrontal est aussi très fortement connecté avec l'aire AIP (Rizzolatti & Luppino, 2001). Les afférences frontales pourraient être à l'origine du contrôle de la sélection des actions qui



peuvent être « générées par soi » (self-generated) ou « induites par le stimulus » (stimulus-driven) selon les intentions de l'agent (Fuster, 2008).



**Figure 5. Le circuit miroir pariéto-frontal.** Vue latérale du cerveau de macaque. Les aires colorées représentent les aires du circuit pariéto-frontal qui contiennent les neurones miroir : le cortex prémoteur ventral (aire F5), l'aire PFG (située entre les aires pariétales PF et PG) et l'aire intrapariétale antérieure (AIP). La scissure intrapariétale (IPS) a été ouverte (jaune clair) pour montrer les aires situées à l'intérieur. Le circuit pariéto-frontal reçoit des informations visuelles de haut niveau des aires du sillon temporal supérieur (STS) et du lobe inféro temporal (IT). Notez que ni l'une ni l'autre de ces aires temporales n'ont de propriété motrice. Le circuit pariéto-frontal est sous le contrôle du lobe frontal, il reçoit des projections des aires motrices pré-supplémentaires (aire F6) et du cortex préfrontal ventral (VPF). Un agrandissement de l'aire F5 présente ses différentes parties. LIP : Aire intrapariétale latérale; VIP : Aire intrapariétale ventrale. Adapté de Rizzolatti, G., & Sinigaglia, C. « The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. » *Nat Rev Neurosci*, (2010)

### 1.2.2. Le système miroir : définition

Le terme « neurone miroir » est de plus en plus abandonné au profit du terme « système miroir » ou « mécanisme miroir » pour caractériser ses propriétés. Le mécanisme général proposé est « *un mécanisme de transformation sensorimotrice qui traduit les informations sensorielles décrivant les actes moteurs réalisés par autrui dans une forme similaire à celle que l'observateur génère lorsqu'il réalise ces mêmes actes* » (Rizzolatti & Fabbri-Destro, 2009). L'intérêt d'une telle définition est qu'elle ne présuppose pas que ce mécanisme soit

réalisé uniquement par des éléments unitaires (neurones), mais que ce soit une propriété qui émerge des interconnexions de plusieurs réseaux neuronaux situés dans des aires aux propriétés différentes. Cela permet d'étendre la notion de « neurones miroirs » dans un réseau plus large d'observation et de compréhension des actions (Cross, Kraemer, Hamilton, Kelley, & Grafton, 2009) et ainsi de pouvoir englober les résultats disparates obtenus chez l'homme.

### **1.2.3. Les études chez l'Homme**

En accord avec les études préliminaires chez l'humain en TEP (Decety, Chaminade, Grèzes, & Meltzoff, 2002; Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Rizzolatti, Fadiga, Matelli, et al., 1996), plusieurs séries d'études en IRMf ont pu mettre en évidence un circuit pariéto-frontal similaire chez l'humain qui posséderait des propriétés « miroir » (Buccino et al., 2001; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Gazzola & Keysers, 2009; Grèzes, Armony, Rowe, & Passingham, 2003; Iacoboni et al., 1999; pour revue et meta-analyse : Molenberghs, Cunnington, & Mattingley, 2012). D'autres études réalisées en EEG et en MEG viennent corroborer les résultats obtenus en imagerie. Les résultats présentent des désynchronisations du rythme enregistré non seulement lors de l'exécution des actions mais aussi pendant l'observation d'actes moteurs réalisés par autrui (Cochin, Barthelemy, Lejeune, Roux, & Martineau, 1998; Hari et al., 1998; Kilner, Marchant, & Frith, 2009; Oberman, Pineda, & Ramachandran, 2007). Par ailleurs, deux études ont notamment attiré notre attention. La première présente des résultats similaires obtenus lors de l'observation de mouvements effectués par des robots en comparaison à ceux d'un humain. Bien qu'il existe des différences dans la forme des deux bras et dans la cinématique des mouvements, le circuit pariéto-frontal était activé dans les deux conditions (Gazzola, Rizzolatti, Wicker, & Keysers, 2007). La deuxième étude est une des rares études qui a réussi à ce jour à démontrer l'existence de mécanismes miroirs au niveau unicellulaire chez l'humain (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010). Tous ces résultats mis en commun permettent de suggérer

que le système observé initialement chez le singe serait aussi présent chez l'homme bien que les outils actuels nous permettent de l'observer le plus souvent de manière indirecte.

Néanmoins, il est à noter que dans la littérature des réseaux miroirs chez l'humain, les mêmes aires déjà observées chez le singe (le gyrus frontal inférieur, le cortex prémoteur ventral, et le lobule pariétal inférieur) sont activés dans des tâches classiques d'observation et d'exécution d'action. En revanche, d'autres aires présentent des propriétés miroirs dès lors que les paradigmes expérimentaux sont constitués de stimuli impliquant d'autres modalités sensorielles. Ainsi, il semblerait que chez l'homme, les propriétés miroirs puissent être étendues au-delà des aires motrices, par exemple Ebisch et al. (2008) ont découvert que l'observation d'une personne touchée par une autre active le cortex somato-sensoriel de la même manière que si l'observateur était touché lui-même, en particulier lorsque l'action de toucher la personne est intentionnelle en comparaison au fait d'être touché accidentellement par un objet (Ebisch et al., 2008). D'autres propriétés similaires ont été découvertes dans des aires auditives, lors de l'écoute d'un son correspondant à une action (Bangert et al., 2006; Ricciardi et al., 2009) ou dans le système limbique pour des réactions émotionnelles (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003). Ainsi il a été proposé que la notion de système miroir pourrait être étendue à une fonctionnalité générale du cerveau. L'idée est notre cerveau va utiliser les mêmes structures impliquées dans le traitement des expériences sensorielles en provenance de notre corps pour appréhender et interpréter ce que nous observons dans le monde et autour de nous (Gallese, 2005).

Toutefois, pour mieux comprendre les raisons qui ont conduit à discuter des implications cognitives de ce mécanisme, il conviendra de revenir sur certaines propriétés fondamentales.

#### **1.2.4. Propriétés et fonctions des neurones miroirs**

Il existe trois distinctions fondamentales dans la littérature des neurones miroirs, qui ont permis de mieux les identifier mais aussi conduit à une certaine « escalade » concernant leur interprétation :

**La correspondance entre action et observation.** De façon générale, les neurones miroirs s'activent lorsque l'animal réalise une action (par exemple porter un aliment à sa bouche). De plus, ils s'activent aussi lorsque l'animal observe des mouvements réalisés par un autre macaque ou un humain. Dans ce cas, l'animal observateur est passif mais ses neurones miroirs vont décharger comme s'il était actif (Rizzolatti, Fadiga, Gallese, et al., 1996). Les neurones miroirs se distinguent des neurones dit « canoniques » qui ont aussi été trouvés dans l'aire F5 et qui s'activent lors de la présentation d'un objet uniquement sans qu'aucune action soit présentée.

Cette première distinction a conduit à préciser que les mécanismes miroirs sont différents du mécanisme d'« affordance » des objets introduit par Gibson (1977). Ce dernier caractérise les propriétés visuelles des objets à pré-activer les différentes « possibilités » d'actions (Gibson, 1977). Ainsi dans une série d'expérience Tucker et Ellis ont pu montrer par exemple que la simple présentation d'objets entraînait chez les participants des effets de compatibilité entre la forme des objets (grands ou petits) et la rapidité de la réponse selon le type de réponse demandé (prise d'une poignée à pleine main ou prise à deux doigts, (Tucker & Ellis, 2004). Il est à noter que pour ces raisons, au cours de nos expériences sur la planification de tâches séquentielles, nous avons pris bien soin de maintenir identique la taille des objets d'une condition à l'autre. Une revue plus détaillée de la littérature sur les effets des propriétés visuelles des objets sur la planification du geste de saisie est proposée en introduction du premier article. Ainsi, si les neurones canoniques vont avoir un rôle évident au cours de la planification motrice du geste de saisie, le rôle des neurones miroirs a tout d'abord été interprété comme un mécanisme de « résonnance » motrice permettant au système moteur de l'observateur de former une représentation motrice de l'action qui serait similaire à celle formée par l'agent lorsqu'il exécute l'action.

**La sélectivité à un acte moteur.** Les neurones miroirs sont sélectifs à l'aspect transitif des actions effectuées. Une action transitive se distingue d'une action intransitive par le fait qu'elle est volontaire et dirigée vers un but. Par ailleurs, le niveau de sélectivité a permis de

distinguer deux catégories, les neurones dit « strictement congruents » correspondent à des neurones qui répondent lorsque l'action est exécutée de manière identique à l'action observée (Gallese et al., 1996). D'autres neurones, dit « grossièrement congruents » sont actifs même si l'action observée est différente de l'action exécutée, comme par exemple un neurone qui serait actif lors de l'exécution d'une action de prise à deux doigts et qui répondrait lors de l'observation de saisie d'un objet à deux doigts ou à pleine main (Rizzolatti, Fogassi, & Gallese, 2001). Ainsi, il a été proposé que ces neurones seraient capables de coder non seulement les actes moteurs mais aussi le but des actions observées.

**La généralisation par le but.** Enfin, un même neurone miroir peut répondre lorsqu'un geste est observé (attraper de la nourriture avec une pince) alors que les mouvements effectués pour réaliser cet acte moteur sont très différents (Umiltà et al., 2008). Par exemple, en entraînant des singes à utiliser des paires de pinces normales ou inverses, dont l'une nécessite d'être serrée alors que l'autre nécessite d'être relâchée pour attraper un objet (Figure 6a), certains neurones qui répondent à l'exécution d'une telle action vont répondre aussi à son observation mais de façon indifférente au type de pince utilisé (Figure 6b). Ainsi, les résultats de Umiltà et collaborateurs montrent que les neurones de F5 répondent de la même façon (pendant la phase de saisie de l'objet avec la pince), bien que les flexions des muscles soient opposées dans les deux situations. Ainsi ces neurones coderaient le but « distal » d'attraper la nourriture et non l'activation des muscles nécessaire pour y arriver (but « proximal »).

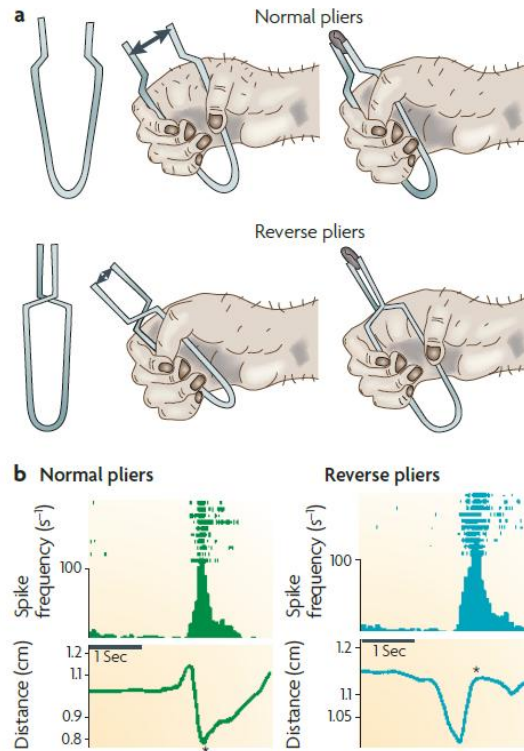


Figure 6. Adaptée de Rizzolatti, G., & Sinigaglia, C. « The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. » *Nat Rev Neurosci*, (2010)

Ainsi la troisième distinction concerne le fait que les neurones miroirs pariétaux et frontaux sont capables d'encoder l'acte moteur observé mais aussi l'action entière dans laquelle l'acte moteur est inclus. Des singes ont été entraînés à attraper des objets avec deux intentions motrices différentes : les placer dans un conteneur ou les porter à leur bouche (Fogassi et al., 2005). Après entraînement, les neurones moteurs de l'aire IPL qui codent pour la saisie (grasping) ont été étudiés selon les deux situations. Les résultats montrent que lors de l'exécution, certains de ces neurones présentent une activité qui varie selon l'action dans laquelle l'acte moteur est inclut ('action-constrained motor neurons'). Les auteurs suggèrent que l'IPL contient des « chaînes » de neurones pour lesquelles chaque neurone pourrait coder un acte moteur tout en étant lié aux autres qui sont sélectifs à un autre acte moteur. Ensemble, ils codent pour une action entière spécifique (par exemple saisir la nourriture pour la porter à la bouche). Ainsi, en plus de décrire ce que l'individu observé est en train de faire (par exemple la saisie), les neurones de l'aire IPL vont aussi aider l'observateur à comprendre pourquoi l'individu réalise cette action, en fonction de la chaîne formée par l'activité de ces

neurones. Autrement dit, selon les auteurs les neurones miroirs de l'aire IPL seraient capables de permettre à l'observateur de *se représenter* l'intention motrice de l'agent. Des études récentes ont montré que des neurones de ce type (action-constrained neurons) étaient aussi présents dans l'aire F5 (Bonini et al., 2010). La comparaison des réponses des neurones miroir de F5 et de IPL (plus spécifiquement dans l'aire PFG) n'a pas permis de montrer de différence significative. Cependant, compte-tenu des connections et du flux important d'information en aller et retour entre ces deux aires, les auteurs proposent que *l'organisation* et la *sélection* des actions intentionnelles se fasse principalement dans le cortex pariétal à travers un accès au « vocabulaire » des différents actes moteurs du cortex pré-moteur et permettre ainsi leur exécution (Bonini et al. 2010, voir aussi : Desmurget & Sirigu, 2009).

#### **1.2.5. Le rôle du système miroir dans la cognition sociale**

Il a donc été proposé que la fonction générale de ce système miroir serait de réaliser une simulation motrice (Hesslow, 2002; Marc Jeannerod, 2001). Cependant, d'autres auteurs ont souhaité aller beaucoup plus loin dans l'interprétation et ont proposé que ce système soit aussi à la base de très nombreuses fonctions cognitives plus complexes. Selon Gallese et Goldman, ce mécanisme serait à la base de la théorie de l'esprit, c'est-à-dire la capacité de se représenter les états mentaux d'autrui en adoptant leur propre perspective (Gallese & Goldman, 1998). Les neurones miroirs pourraient permettre ainsi de comprendre les actions et les intentions de l'agent et serait primordiaux dans l'émergence du comportement d'imitation (Iacoboni et al., 2005). De manière générale, les liens directs entre la perception et l'action seraient les briques constitutives de la compréhension et des interactions sociales (Knoblich & Sebanz, 2006).

Certains auteurs vont encore plus loin en proposant que le mécanisme miroir et l'imitation seraient la force conductrice derrière le « grand saut en avant » dans l'évolution de l'Homme (Ramachandran, 2000). Bien que spéculatives, ses réflexions vont jusqu'à suggérer que ces neurones permettent de créer des représentations de second-ordre ou meta-représentations des

processus du cerveau. Il s'agirait alors des bases neurales de l'introspection et de la réciprocité entre conscience de soi et conscience des autres (Ramachandran, 2000).

Pour Jean Decety et Thierry Chaminade, le cortex pariétal et le cortex préfrontal sont les bases neurophysiologiques de la distinction entre soi et les autres en lien avec le phénomène d'agentivité (le fait d'être conscient de ses propres actions). Les états mentaux, privés par essence pourraient alors être partagés entre les individus de sorte d'obtenir une *représentation partagée* d'un même événement. Par exemple, en présentant à des participants des visages représentant une émotion pendant la lecture d'une histoire congruente ou non avec l'expression présentée, ils ont pu montrer que l'écoute d'histoires tristes entraînait des réponses plus fortes dans des structures reliées au traitement des émotions (amygdale et lobe temporal adjacent) par rapport à l'écoute d'histoires neutres (Decety & Chaminade, 2003a). De plus ils ont montré que les régions du cortex prémoteur, du lobule pariétal, du pre-SMA et de la scissure centrale étaient aussi associées au fait d'écouter une histoire triste. Ils suggèrent alors que ce réseau est recruté pour simuler les expériences affectives des protagonistes. Bien que conscient que d'autres structures sont aussi impliquées dans le traitement des émotions (Decety & Chaminade, 2003b), ces résultats ont permis aux auteurs de suggérer que ces réseaux extraient les intentions du contenu narratif, pour pouvoir les interpréter et ainsi générer chez les participants un sentiment de compassion avec les visages présentés, principalement lorsque leur expression est congruente avec le récit.

Ainsi, chez l'humain comme souligné ci-dessus, le système miroir qui dépasse le cadre des simples représentations motrices permettrait de relier les liens directs entre action et observation avec les capacités d'empathie (Decety & Jackson, 2004; Gallese, 2001; Keysers, Kaas, & Gazzola, 2010). Cependant, ce rôle particulier du système miroir est largement critiqué et est encore aujourd'hui un grand sujet de débat. Avant d'aller plus loin, il convient de s'arrêter quelques instants sur les limites des modèles basés sur la correspondance directe entre action et observation.



### **1.3. Limites et critiques des Neurones Miroirs**

Comme présenté brièvement ci-dessus, des auteurs soulignent l'importance du système miroir à la fois pour l'apprentissage par imitation et comme base pour des capacités cognitives de plus haut niveau. Mais ces propositions sont largement discutées encore aujourd'hui et de nombreux travaux expérimentaux sont encore à ce jour nécessaires pour en explorer les limites. Voici quelques une des critiques les plus intéressantes qui permettent de mettre en évidence les limites du système miroir.

#### **1.3.1. Des résultats contradictoires chez l'homme.**

Jusqu'à très récemment (Mukamel et al., 2010), il n'existait pas à ce jour d'étude qui avait réussi à démontrer l'existence de neurones aux propriétés miroirs dans le cortex humain et beaucoup de chercheurs se réfèrent à des cellules « supposées » (putatives) lorsqu'ils discutent des neurones miroirs chez l'humain. En revanche, plusieurs études ont mis en exergue certains points discutables chez l'humain (voir (Hickok, 2009)). Par exemple, Lingnau et collaborateurs (2009) ont montré en utilisant un paradigme intéressant qu'il pouvait y avoir des asymétries dans le phénomène d'adaptation des aires cérébrales qui contiendrait ces neurones. Plus précisément, le phénomène d'adaptation est le mécanisme qui réduit la réponse des neurones lors de la présentation du même stimulus de manière répétée. Ils ont trouvé que lorsque les participants observaient des actes moteurs puis les exécutaient, cette adaptation était présente. En revanche, le cas opposé, exécution suivi d'observation répétée, ne montrait pas le même phénomène d'adaptation (Lingnau, Gesierich, & Caramazza, 2009). Bien que d'autres auteurs ont trouvé les résultats opposés, c'est à dire une adaptation uniquement lorsque les participants exécutent d'abord et observent de façon répétée ensuite (Chong, Cunnington, Williams, & Mattingley, 2009), ces auteurs en ont conclu que ces asymétries d'adaptation des réponses neuronales étaient incompatibles avec l'hypothèse d'une correspondance directe et bidirectionnelle entre observation et action. Ils suggèrent alors que les réponses dans les aires du système miroir pourraient refléter une facilitation motrice (Brass, Bekkering, & Prinz, 2001; Fadiga et al., 1995). Autrement dit, que l'observation d'un

mouvement aurait une influence automatique sur son exécution (Brass, Bekkering, Wohlschläger, & Prinz, 2000). Pour d'autres auteurs, un des problèmes de la nomenclature des neurones miroirs et qu'elle prend en compte trop de cas disparates, ils posent même la question en ces termes : « Quand pouvons-nous nous arrêter de les appeler neurones miroir ? » (Uithol et al. 2008). Ils précisent par exemple que si la définition correspond bien aux neurones strictement congruents et qu'elle pourrait s'étendre à des neurones grossièrement congruents, il existe plusieurs variétés de congruence, dépendant du niveau d'abstraction auquel on se réfère. Par exemple, il existe aussi des neurones miroirs « non congruents », qui ne répondent pas à la même action exécutée qu'à celle observée. Pour ces neurones la correspondance se ferait alors seulement au niveau le plus abstrait, celui de la correspondance entre deux actions « orientées vers un but » (Uithol et al., 2008). Tous ces éléments soulignent la nécessité des définitions en particulier au regard des niveaux d'abstraction les plus élevés et bien que ces résultats ont permis des avancées majeures dans la compréhension du fonctionnement cognitif, des révisions seraient nécessaires dans les modèles et leur interprétation afin de préserver l'équilibre entre théorie et résultats empiriques (Mahon & Caramazza, 2008).

### **1.3.2. Les théories motrices de la cognition sociale : un débat éclairé**

Un des points clé du débat a récemment été l'objet d'une controverse (Jacob & Jeannerod, 2005), à propos du rôle des neurones miroirs dans la compréhension des actions et surtout son extension, la compréhension des intentions de l'agent. Jacob et Jeannerod présentent alors une controverse sur l'hypothèse d'un mécanisme unique pour les interactions motrices et les interactions sociales :

« Considérez le Dr Jeckyl et Mr Hyde. Le premier est un chirurgien de renommée qui réalise des appendicectomies sur ses patients anesthésiés. Le second est un dangereux sadique qui réalise exactement les mêmes mouvements de la main sur des victimes non anesthésiées. Ces deux êtres sont en fait les deux personnalités d'un seul et même individu. Supposez

maintenant que le Dr Watson soit le témoin d'un acte réalisé par cet individu sans voir la personne sur laquelle cet acte est réalisé. Le Dr Watson va donc parfaitement percevoir la façon dont il prend son scalpel et applique la section sur la partie du corps. L'intention motrice du Dr Jeckyl et de Mr Hyde est exactement la même dans les deux situations, cependant leur intention sociale est très différente. Alors que l'un cherche à soigner son patient, l'autre cherche à tirer du plaisir de l'agonie de sa victime ». Jacob & Jeannerod en ont conclu qu'en observant uniquement les mouvements de l'agent (celui qui réalise l'action), l'observateur va être capable de se représenter l'intention motrice, mais pas son intention sociale.

Selon ces auteurs, la cognition humaine comprend tous les processus nécessaires à la perception et à la compréhension d'autrui. Donc elle inclut, mais elle ne peut pas se résumer aux seuls processus impliqués dans la compréhension des actions d'autrui. Plus précisément, toutes les actions humaines ne sont pas dirigées vers des cibles inanimées et certaines sont dirigées vers des individus. Au-delà de la distinction entre les intentions motrices et les intentions préalables à l'action (voir (Searle, 1983), les intentions non sociales d'un agent doivent pouvoir être distinguées de ses intentions sociales, c'est à dire les intentions d'agir en direction d'un individu qui à la différence d'un objet inanimé va agir aussi en retour. Dès lors, une intention sociale devient l'intention d'affecter le comportement d'autrui. Et parce que les humains agissent à travers des représentations mentales, une intention sociale serait alors l'intention de modifier la représentation mentale d'autrui. La question est posée : « Est-ce qu'un observateur peut se représenter l'intention sociale d'un agent simplement en observant ses actions ? ». Les auteurs suggèrent que la capacité d'un observateur à faire correspondre les mouvements de l'agent dans son propre répertoire moteur n'est pas forcément la même capacité que celle de se représenter ses intentions sociales et que d'autres structures cérébrales en seraient à l'origine.

Pour d'autres auteurs, la distinction ne doit pas être aussi stricte et qu'il n'y a pas encore de preuves suffisantes pour distinguer le système miroir d'un système de "cognition sociale" à

proprement parler (Pierno & Ansuini, 2007). Pour d'autres auteurs certains systèmes de "bas niveau" pourraient suffire pour que deux personnes puissent se coordonner, notamment dans des actions conjointes simples (Vesper, Butterfill, Knoblich, & Sebanz, 2010). Ce large débat a conduit à un nouvel intérêt dans les sciences cognitives et en psychologie expérimentale pour les liens entre cinématique et interactions sociales (Becchio, Manera, Sartori, Cavallo, & Castiello, 2012; Ferri, Campione, Dalla Volta, Gianelli, & Gentilucci, 2011). Bien que particulièrement inspirant et pertinent pour mieux comprendre les situations interactives comme le cas des actions conjointes (Sebanz, Bekkering, & Knoblich, 2006), ces résultats ne trouvent pas forcément un écho direct dans le champs des modèles computationnels de l'action. Pourtant, il nous semble particulièrement important de mieux comprendre ce type de phénomènes dès lors que l'on s'intéresse à la création d'une architecture robotique d'assistance.

Au cours des travaux expérimentaux que nous avons réalisés, nous allons chercher à démontrer deux hypothèses principales. La première hypothèse est l'idée qu'il existe un continuum entre les actions individuelles et les actions dites « sociales ». Ainsi, les modèles utilisés par le cerveau pour gérer les situations d'interaction avec l'environnement pourrait pour certains être directement transposables dans le cadre des interactions sociales. En revanche, il existerait des spécificités reliés uniquement aux interactions sociales, notamment vis à vis des intentions d'*affecter le comportement d'autrui*. La deuxième hypothèse est qu'il existe dans les mouvements que nous effectuons des indices extérieurs, cinématiques, permettant de lire non seulement les intentions *motrices* mais aussi les intentions *sociales* et que nous sommes suffisamment sensibles à ce type d'indices pour pouvoir distinguer et catégoriser les intentions *sociales* d'un mouvement dont l'objectif moteur proximal serait identique.

# **Chapitre 2. Modèles computationnels de l'action et approches méthodologiques**

Pour continuer à élargir le cadre dans lequel nous avons effectué nos recherches, au cours de ce second chapitre nous allons nous intéresser à la deuxième grande fonction impliquée comme responsable du phénomène de la “uncanny valley”, c’est à dire les capacités prédictrices du système moteur humain et sa modélisation computationnelle dans le cadre des modèles internes. Pour éviter d’entrer dans une revue complète et fastidieuse de tous les modèles existants, nous avons choisi de nous concentrer dans cette introduction uniquement sur la théorie des modèles internes, car contrairement à d’autres modèles, leur pouvoir explicatif dépasse largement le cadre de quelques effets psychométriques. De plus, ils ont pu être étendu à des situations d’interaction sociales et sont largement réutilisés pour expliquer des résultats en psychologie et en sciences cognitives.

## **2.1. Les modèles internes**

Depuis plusieurs années, le concept de “modèle interne” s’est largement développé dans la littérature du contrôle moteur (Jordan, 1996; Kawato, 1999; Miall & Wolpert, 1996). Ce concept s’appuie sur l’hypothèse que le cerveau dispose d’une connaissance intrinsèque des mécanismes sous-jacents au contrôle moteur. Ceci pourrait résulter d’apprentissages sur le long terme mettant en relation les systèmes sensoriels, neuronaux et musculaires. Ainsi, le système nerveux serait capable de reproduire en “interne” les caractéristiques d’entrée/sortie de l’appareil moteur humain dans le but de planifier et de contrôler le mouvement (Kawato, 1999). Il existe deux types de modèles internes, le modèle inverse et le modèle direct (forward). Le modèle direct prédit les conséquences sensorielles à partir des copies d’efférence issues directement des commandes motrices (ou décharges corolaires). Le modèle

inverse va de manière opposée, calculer les commandes motrices nécessaires pour obtenir une trajectoire désirée à partir par exemple de la position d'une cible à atteindre.

### **2.1.1. Mise en évidence des modèles internes**

L'existence de tels modèles a été démontrée par plusieurs études au cours desquels des participants devaient simplement pointer en direction de cibles sur l'écran alors que leurs bras étaient sous influence d'un champ de force. Ces champs de forces pouvaient être appliqués par des manipulandum robotiques (Shadmehr & Mussa-Ivaldi, 1994) ou des chambres rotatives (Lackner & Dizio, 1994). Les auteurs ont tout d'abord trouvé que de larges adaptations se faisaient extrêmement rapidement et que les sujets pouvaient en quelques essais réaliser des trajectoires rectilignes à l'écran alors que leur bras décrivait des courbes importantes. Ce qui était encore plus intéressant est le fait qu'en supprimant le champ de force, et sans retour visuel de leur propre membre, les participants conservaient ces adaptations et des erreurs terminales importantes apparaissaient dans la direction opposée à celle du champ de force préalablement imposé. Ainsi, de telles adaptations et leur maintien parfois très longtemps après la fin du champ de force (250 essais), semble indiquer que les modèles de dynamique inverse du bras sont stockés au niveau cérébral et réutilisés ensuite dans le cas où par exemple, la vision ne peut pas aider à guider le bras.

### **2.1.2. Modélisation computationnelle**

Dans le cadre des systèmes dynamiques, l'objectif de ce type de modèle est de décrire les changements d'états du système. Par exemple, le modèle direct peut prédire l'état suivant à partir de l'état courant et de la commande motrice en cours. Il pourra alors également estimer la sortie sensorielle correspondante à la réalisation finale de cette commande. En revanche, les modèles inverses cherchent à remonter aux causes qui ont provoqué la modification de l'état du système, et donc la variation des informations sensorielles enregistrées entre les deux états. Ainsi, à partir de la situation d'arrivée, l'objectif est de donner une estimation des commandes nécessaires pour aboutir à la variation désirée. Ce type

de modèle est particulièrement intéressant pour déplacer un bras de robot en connaissant sa position initiale et la position d'arrivée. Par ailleurs, Miall et Wolpert (1996) séparent en deux niveaux le modèle direct, un premier niveau appelé modèle dynamique direct qui se situe au niveau musculaire, et qui va *effectuer les changements d'états* à partir des états précédents et des commandes motrices, puis un deuxième niveau appelé modèle sensoriel direct, qui à partir de l'état déterminé par le modèle dynamique, va *prédire les conséquences sensorielles* et ainsi terminer la boucle pour revenir au système nerveux central (Figure 7).

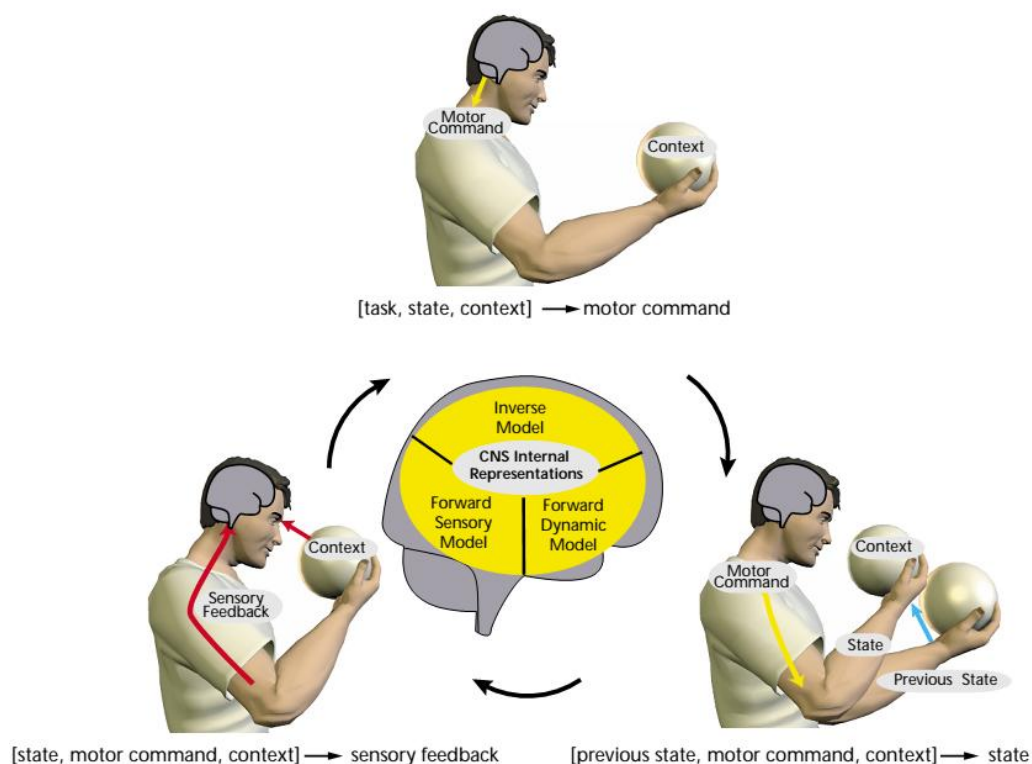


Figure 7. Adaptée de Wolpert, D. M., & Ghahramani, Z. Computational principles of movement neuroscience. *Nature Neuroscience* (2000).

D'après ces auteurs, cette boucle sensori-motrice en trois étapes peut servir de base pour modéliser à la fois 1) comment les tâches conditionnent les comportements, 2) comment les commandes motrices sont générés, 3) comment les états ou les contextes peuvent être estimés et prédits et 4) comment ces modèles internes seraient représentés et appris chez l'homme (Wolpert & Ghahramani, 2000). Tout ceci mis en commun, nous comprenons facilement pourquoi ce type de modèle est très utile pour être implémenté dans des architectures

artificielles ou en réalité virtuelle. De plus, ce type de modèle permet de comprendre l'importance du corps en tant que substrat des apprentissages et en ça il pourrait permettre de relier les théories incarnées de la cognition avec les questions de robotique développementales.

Par ailleurs, un autre des grands intérêts de ce type de modèle est qu'il permet de faire des comparaisons directes entre les conséquences sensorielles obtenues et les conséquences réelles nées de l'interaction de l'agent avec son environnement. Dès lors, un signal d'erreur de prédiction est généré pour éventuellement effectuer des changements de modèle si quelque chose d'imprévu a modifié les conditions initiales, ou si des perturbations non modélisées ont agi sur le système. Néanmoins, cela nécessite qu'il existe un grand nombre de modèles internes disponibles rapidement pour le cerveau. Il a alors été proposé que le cervelet, d'après sa configuration anatomique et ses connexions nombreuses avec le cortex, soit la structure qui encoderait cette multitude de modèles nés des apprentissages moteurs (Kawato, 1999; Kawato et al., 2003; Wolpert, Miall, & Kawato, 1998). Par exemple, le modèle MOSAIC (Wolpert & Kawato, 1998) propose une architecture de ce type et plusieurs extensions de ce modèle ont depuis été proposées (Davidson & Wolpert, 2004; Haruno, Wolpert, & Kawato, 2001, 2003) et discutées (pour revue voir : Davidson & Wolpert, 2005).

### **2.1.3. Un cadre théorique commun pour le contrôle moteur et les interactions sociales**

La théorie des modèles internes a donné lieu à des propositions intéressantes concernant un cadre computationnel pour les interactions sociales. Par exemple, si le contrôle moteur est intéressé par les relations d'entrées/sorties pour contrôler un objet du corps (comme le bras), il est possible de faire une analogie avec une situation impliquant deux personnes en interaction. Dans ce cas, l'objet que l'on cherche à modifier est le *comportement d'autrui* à travers des signaux qui ne sont pas internes mais externes (Figure 8). Les



commandes motrices deviennent alors des signaux de communication ou des attitudes du corps. Lorsqu'ils sont perçus par le partenaire, ils peuvent avoir une influence sur ses états mentaux. Ces états mentaux vont constituer un ensemble de paramètres définissant son *comportement*. Nous pouvons alors considérer le partenaire comme une personne qui va passer par une série d'états comme c'est le cas pour notre propre corps. Si nous connaissons l'état actuel de notre partenaire et que nous possédons un modèle de son comportement, il nous est alors possible d'anticiper la réponse qu'il va donner selon ce que nous allons lui dire ou montrer. Il nous est ensuite possible de percevoir les conséquences de nos propres actions sur le comportement du partenaire (Wolpert, Doya, & Kawato, 2003). Ce signal perçu en retour (ou *feedback*) pourra ensuite nous servir comme un signal d'erreur si notre objectif premier n'a pas été atteint. Ainsi, dans le cadre des interactions sociales, nous serions capables de donner une estimation des états mentaux internes d'autrui mais aussi de former de nouveaux modèles inverses ou directs du comportement du partenaire, après apprentissage à travers les interactions que nous avons avec chaque individu.

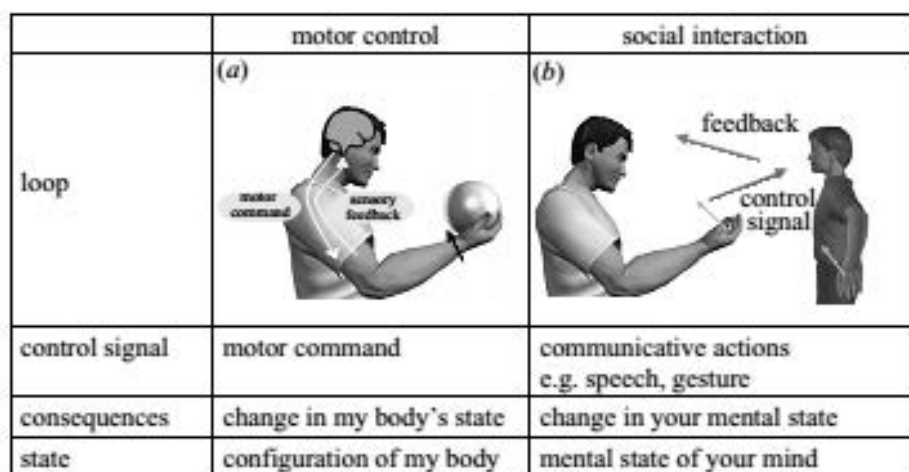


Figure 8. Adaptée de Wolpert, D. M., Doya, K., & Kawato, M. A unifying computational framework for motor control and social interaction. *Phil. Trans. of the Royal Society of London. Series B: Biological Sciences.* (2003).

Ce type de modèle est particulièrement intéressant du point de vue computationnel et a donné lieu à des avancées significatives dans le domaine de l'inférence des états mentaux (ou théorie

de l'esprit) en robotique (Breazeal, Buchsbaum, Gray, Gatenby, & Blumberg, 2005; Breazeal, Gray, & Berlin, 2009; Oztop, Wolpert, & Kawato, 2005; Scassellati, 2002).

Toutefois, dans ce cadre, le problème à résoudre reste toujours très complexe. Par exemple, le cerveau doit composer avec énormément de bruit parmi les signaux sensoriels et moteurs, et parfois aussi avec des délais importants notamment alors que certaines tâches doivent être réalisées extrêmement rapidement. De plus, pour réaliser une certaine tâche le cerveau doit trouver une solution parmi une énorme redondance dans ses effecteurs. En d'autres termes, il existe une infinité de possibilités pour atteindre un même but et pourtant les humains exécutent des mouvements relativement similaires lorsqu'on leur demande de répéter plusieurs fois la même tâche. Ce dernier problème appelé le problème des "degrés de liberté" a été proposé initialement par Bernstein (1967) et de très nombreuses propositions au niveau computationnel ont été faites pour le résoudre. Parmi les différentes classes de modèles, une classe a particulièrement attiré notre attention, il s'agit des modèles dit "optimaux" dont l'intérêt principal réside dans la capacité à transformer un *critère de performance* en prédictions concernant le comportement du système (Todorov, 2004).

#### **2.1.4. Des modèles d'optimisation pour le contrôle sensori-moteur**

Il existe de très nombreux modèles de contrôle optimaux du mouvement. La plupart d'entre eux sont basés sur un critère de performance simple (variance, énergie, précision). Ce critère appelé « coût » doit être minimisé dans le but de générer des trajectoires motrices qui atteignent l'objectif de la tâche, tout en respectant un certain nombre de lois biologiques. Ces lois dites « biologiques » sont multiples. Ainsi le trade-off entre vitesse et précision (Fitts, 1954), le principe d'isochronie (P. Viviani & Schneider, 1991), les relations linéaires de vitesse articulaire (Soechting, Lacquaniti, & Terzuolo, 1986), les profils de vitesse en cloche pour les mouvements d'atteinte (M. Jeannerod, 1984, 1988) ou la relation inverse entre la vitesse instantanée et la courbure de la trajectoire appelée loi de puissance deux-tiers (Lacquaniti, Terzuolo, & Viviani, 1983) sont autant d'observations démontrées chez l'homme qui cherchent

à être reproduites par le modèle. L'idée principale est donc qu'un seul critère pourrait permettre de résumer les différentes observations des invariants du mouvements humain. Ainsi, plusieurs propositions ont été faites comme l'énergie consommée par les muscles dans le cadre de la locomotion (Davy & Audu, 1987; Popovic, Stein, Namik Oguztoreli, Lebedowska, & Jonic, 1999), la dérivée de l'accélération (*minimum-jerk* : Flash & Hogan, 1985; Todorov & Jordan, 2002), la variation de couple articulaire (Uno, Kawato, & Suzuki, 1989) ou la précision terminale (Harris & Wolpert, 1998). Ainsi, les modèles optimaux ont été parmi les plus efficaces pour prédire de très nombreux résultats empiriques chez l'humain. Dès lors, les théories optimales du contrôle moteur ont permis de mieux comprendre pourquoi les humains agissent de façon quasi régulière. Néanmoins, il semblerait qu'en réalité, plusieurs critères d'optimalité soit utilisés de manière combinés entre eux (D. A. Rosenbaum, Slotta, Vaughan, & Plamondon, 1991).

Plus récemment, les auteurs ont proposé d'optimiser le contrôleur en ligne, (*optimal feedback controler*) qui va, à chaque étape de l'exécution, vérifier si la trajectoire actuelle est conforme avec la trajectoire prédite. Si l'effecteur s'en écarte, il faut alors réaliser les corrections nécessaires. Dans le cas des modèles présentés ci-dessus, l'optimisation se fait généralement sur la "boucle ouverte" en générant de nouvelles commandes. Le contrôle en ligne est effectué par un asservissement basique qui va simplement annuler les déviations observées. En revanche, dans un modèle de contrôle qui optimise aussi la "boucle fermée", le contrôle en ligne est entièrement programmable, c'est à dire qu'il va construire et sélectionner les meilleures transformations possibles entre les états de l'appareil ou de l'environnement en commandes motrices (Guigon, Baraduc, & Desmurget, 2003, 2008; Todorov & Jordan, 2002). Ainsi, plutôt que de se baser sur des idées préconçues concernant les commandes à appliquer en cas de déviation, le contrôleur va laisser l'effecteur, la tâche, ou les contraintes guider le schéma de contrôle qui correspond le mieux à ce qui doit être réalisé. L'expérience présentée dans le premier article du Chapitre 3 expose des résultats en faveur de l'hypothèse selon laquelle il existerait effectivement plusieurs modes de contrôle qui varient selon les

contraintes appliquées dans la tâche. Cela aura aussi pour conséquence la prise en compte de l'ensemble ou d'un seul des éléments de la séquence motrice, préalablement à son exécution.

Enfin, si ces modèles de contrôle permettent de mieux comprendre comment les trajectoires sont générées et apprises dans des systèmes individualisés, il devient beaucoup plus difficile d'appliquer ces modèles dans des cadres mêlant contrôle optimal et interactions sociales. Une des critiques majeurs de ce type de modèle est qu'il se base essentiellement sur la théorie des modèles internes dont la démonstration n'est pas évidente (Friston, Mattout, & Kilner, 2011). En particulier, l'hypothèse des modèles internes suggère que les mêmes processus de prédiction motrice soient utilisés à la fois pour l'exécution des actions mais aussi pour leur observation. Cependant, il est possible que d'autres capacités d'inférences soit utilisées en particulier dans le cas des observations d'autrui (Friston et al., 2011), alors que ce ne serait pas le cas lors de la simple programmation de gestes dans un environnement statique. Nous reviendrons sur ce point dans la discussion à la lumière des résultats expérimentaux qui sont exposés dans les chapitres suivants dans ce manuscrit.

## **2.2. Methodologie expérimentale**

C'est dans ce cadre théorique à la fois neuronal, cognitif et computationnel que s'inscrivent les recherches du projet INTERACT, dont l'objectif est de créer un robot humanoïde qui aurait la capacité par observation et imitation de réaliser des tâches conjointes avec l'homme. Pour y parvenir, les chercheurs de l'ETIS (Univ. Cergy-Pontoise) et du LISV (Université de Versailles) se sont associés avec des chercheurs en psychologie et sciences cognitives l'URECA (Université Lille 3) pour s'intéresser particulièrement à la manière dont les humains au cours des interactions sociales sont capables de lire et prédire les intentions d'autrui. Car en effet, doter les robots de capacités de lecture d'intentionnalité est un des challenges importants de le domaine de la robotique d'assistance pour améliorer la sécurité et l'intuitivité des interactions non-verbales avec l'humain.

Nous avons donc conduit au total cinq expériences en comportement moteur qui sont décrites en détail dans les chapitres suivants. Cependant, il est à noter que notre méthodologie s'inspire du challenge que représente la lecture d'intention sociale proposé par Jacob et Jeannerod (2005). Il nous a donc semblé important de récolter des données très précises en cinématique chez l'humain pour comprendre comment les séquences motrices impliquant des manipulation d'objets étaient planifiées. D'autre part, un accent important a été mis sur les aspects écologiques des tâches que nous avons mis en place ainsi que sur les enregistrements en temps réel afin de nous assurer que nos mesures puissent refléter le comportement des personnes dans des situations naturelles.

L'objectif des deux premières séries d'expériences (article 1 et 2) a été de connaître les caractéristiques détaillées de la cinématique du mouvement d'une tâche séquentielle impliquant deux éléments moteurs : une phase d'atteinte et de saisie suivi d'une phase de déplacement et de pose de l'objet sur le réceptacle. Des modifications des contraintes imposées, des instructions ou des caractéristiques du réceptacle terminal ont été proposées pour vérifier leurs influences sur les différentes phases de la séquence. Pour les participants, la situation leur paraissait « interactive » et « agréable » (évalué sur une échelle analogique) dans le cas où les conditions étaient en temps réel et en absence d'instruction verbales (Voir Chapitre 3.1).

En parallèle, dans une série d'étude en observation (article 3 et 4), nous sommes parti d'une situation d'interaction compétitive sous forme de jeu (« Jungle Speed »), au cours duquel l'objectif était d'attraper l'objet au centre de la table le plus rapidement possible. Néanmoins, la séquence d'action correspondant à la compétition ne fut pas utilisée car trop différente d'une situation d'interaction réelle généralement réalisée à vitesse naturelle (« préférée »). Nous nous sommes donc spécifiquement intéressé aux phases de préparation du jeu (mise en place) et de récompense selon la personne qui a « gagné » le tour de jeu. Durant les phases de préparation ou de récompense, nous avons pu enregistrer des interactions à vitesse préférée entre les participants. Bien qu'elles étaient évidemment soumises aux règles

du jeu (mettre l'objet en position de départ avant chaque essai ou donner l'objet à la personne qui a gagné le round précédent), et malgré la présence évidente des caméras, les deux participants n'étaient pas du tout au courant que ces phases-là du jeu seraient celles qui seraient ensuite conservées comme stimuli. De façon intéressante, les enregistrements cinématiques de ces séquences ont présenté des caractéristiques différentes selon les étapes du jeu et du but final de l'action (voir Chapitre 4.1 partie A : Creating the Stimuli), bien que dans toutes ces situations, les mouvements d'atteinte et de saisie de l'objet se faisaient à partir de la même position de départ, et pour une même position de l'objet à atteindre.

Cependant, nous n'étions pas suffisamment convaincu de répondre entièrement au paradigme théorique de Jacob et Jeannerod, en particulier parce que les positions terminales de la séquence variaient fortement d'une phase de jeu à l'autre (préparation vs récompense) et d'une catégorie de séquence à l'autre (donner « à toi » vs prendre « pour moi »). Nous nous sommes donc inspirés de cette situation pour créer un nouveau jeu, dans lequel la première étape du jeu était rigoureusement identique dans toutes les conditions (Chapitre 5), que ce soit pour la prise de l'objet et pour sa pose, lors de la mise en place du jeu (« Preparatory Action »). Nous avons ensuite pu manipuler les caractéristiques du jeu (prédictibilité), la présence et la position du partenaire ou le contexte *d'intention sociale* dans lequel cette action de préparation était réalisée. Ainsi, nous avons pu proposer une vérification expérimentale du paradigme théorique de Jacob et Jeannerod et cela nous a permis de faire des constatations très intéressantes sur les différentes *composantes* de l'interaction sociale, apportant ainsi un éclairage nouveau sur les liens entre cinématique, intentions et contexte social des actions.





## **Chapitre 3. Exécution**



# **3.1. Coupled Planning of sequential Pick and Place under Speed and Accuracy constraints**

Abstract :

Planning a sequence of two motor elements is much more than concatenating two independent movements however, very little is known about the cognitive processes that are needed to perform fluent sequences for intentional object manipulation. In this series of studies, the subjects' task was to reach for, and pick to place a wooden cylinder to set it on a place pad of 3 different diameters, which served to modify terminal accuracy constraints. Subjects were required to perform the sequences (1) at their preferred speed or (2) as fast as possible. Action kinematics was recorded with the Qualysis motion-capture system in order to implement a real-time protocol to engage subjects in a true interactive relation, and to encourage them to comply with the experimental constraints. Results revealed that with low internal constraints (at preferred speed), low coupling between the 2 elements of the motor sequence was observed, suggesting a step-by-step planning strategy. Under high constraints (at fastest speed), an important terminal accuracy effect back propagated to modify early kinematic parameters of the first element, suggesting strong coupling of the parameters in an encapsulated planning strategy. In studies 2 and 3, we further manipulated instructions and timing constraints to confirm the importance of time and context predictability for coupled planning. These findings overall sustain the hypothesis that motor intention can be read through the detection of early kinematic modulations and suggest that global planning may be the strategy required for intuitive social interaction both in human-human and in human-robot interplay.

**Theme:** Action planning

**Keywords:** Kinematics, motor planning, sequential movement, grasping, speed-accuracy trade-off.

## **Introduction**

Our ability to interact with others depends mostly on our capacity to anticipate motor goals and as such researchers have asked the question of how the brain is capable of anticipating the intention of a partner through simple observation of body movements. One of the most interesting findings of the last decade has been the mirror system which is known as “*a mechanism that translates sensory information describing motor acts done by others into a motor format similar to that the observers themselves generate when they perform those acts*” (Rizzolatti & Fabbri-Destro, 2009). Hence, through the use of the mirror system, individuals may gain access to the understanding of motor goals through the direct visual observation of the kinematic properties of the actors’ movements. Recently, experimental data has confirmed this working hypothesis by showing that the simple observation of motor kinematics provides key information for the understanding and the anticipation of motor actions performed by conspecifics during social interaction (Becchio et al., 2012; Campione & Gentilucci, 2011). A large series of psychological studies have further shown that it is the slight kinematic variations in the early moments of the action that seems to cue the reading of motor intention (Manera, Becchio, Cavallo, Sartori, & Castiello, 2011; Manera, Schouten, Becchio, Bara, & Verfaillie, 2010; Sartori, Becchio, & Castiello, 2011). Nevertheless, these findings do not explain why those kinematic variations could be seen as early as in the first reach-to-grasp movement. The aim of the present series of experiments was to assess such question for a simple ecological action that required the sequencing of multiple motor elements as it is the case in a pick and place task that is directed towards a small physical object,

The question of intention reading is important for the scientific community, and is directly related to the better understanding of functional properties of movement kinematics, the

question being whether there is a direct relationship between contextual effects and action planning laws (Ansuini, Santello, Massaccesi, & Castiello, 2006; Ansuini, Giosa, Turella, Altoè, & Castiello, 2008; Becchio, Sartori, & Castiello, 2010; Ferri, Campione, Dalla Volta, Gianelli, & Gentilucci, 2011). How the environment in which an action is produced will modify the planning properties of that movement is a key point notably in social robotics, a new field that now needs more accurate theoretical models of motor psychology in order to instrument intuitive robot-human interactions (Lewkowicz, Delevoye-Turrell, Bailly, Andry, & Gaussier, 2013; Quesque, Lewkowicz, Delevoye-Turrell, & Coello, 2013). Recent studies have suggested indeed that humanoid robots need to move following the biological properties of human movements in order to enable intuitive interaction (Chaminade & Cheng, 2009; Chaminade, Franklin, Oztop, & Cheng, 2005). The findings presented here will give valuable information on how contextual constraints will specifically affect the kinematics of sequential motor actions during pick and place movements and hence, will offer important cues on what body information is required for the cognitive ability of intention reading. The transferability of this empirical work can be as wide spread as the construction of artificial neural networks for humanoid robots (needing *intuitive interaction*) to the development of clinical neuropsychological tools (searching for guidelines in the *rehabilitation of abnormal social skills*).

In the following, we will provide a brief description of the literature describing the effects of speed and accuracy constraints on arm kinematics for pointing movements. We will then turn to the question of contextual effects in object manipulative tasks that are less frequently used in experimental psychology paradigms.

## **Sequential pointing**

**One Target Advantage.** According to Adam and colleagues (2000), movements embedded in a sequence are not independent but may mutually influence each other. The one

target advantage model was suggested after the finding that a rapid aiming movement is executed faster when it is allowed to stop on the target than when it must proceed and hit a second target (Adam, van der Bruggen, & Bekkering, 1993). Adam and colleagues then proposed the *target impact constraints hypothesis*: The one-target advantage is the consequence of a motor control organization in which one-element aiming responses exploit future impacts with targets as a passive control mechanism to decelerate and optimize limb control. This one-target advantage would reflect inter-influences among motor segments (Adam et al., 2000) and, as such, would belong to a general class of phenomena called *context effects*.

**Segment interdependency and Fitts' Law.** If those movements embedded in a sequence are not completely *independent*, one can name them *inter-dependent*, which is the phenomena described by Rand and colleagues in a series of experiments involving rapid two-segment aiming movements. By varying the index of difficulty (ID) of the first or the second target, the authors reported context-dependent kinematic changes in the performance of the initial segment. Movement duration increased when ID was increased through the reduction of the second target size (Rand, Alberts, Stelmach, & Bloedel, 1997). These later results were later replicated by Weiss and colleagues in elderly controls and patients with Parkinson's disease (P. Weiss, Stelmach, & Hefter, 1997). More specifically, they found that in both groups, the movement time of the first segment was prolonged in the more difficult condition, i.e., 'small' vs. 'large' size of the second target. More specifically, the amplitudes of peak velocity, peak acceleration and peak deceleration were reduced under increased accuracy requirements in both groups, thus leading to a lengthening of both times to peak velocity and peak deceleration. In a second experiment, Rand and colleagues asked whether the interdependence kinematic changes were diminished when difficulty of the first segment was increased. They found that when the ID was low in the first segment, there was a back propagation effect of the second target ID on the first motor segment: movement duration increased with a lower peak velocity. In contrast, when the ID was high in the first segment,

the interdependencies disappeared: movement duration and peak velocity of the first segment were unaffected by the difficulty of the second segment (Rand & Stelmach, 2000). Taken together these results strongly suggest that for pointing movements, when planning a sequence of two motor elements, the constraints of the second segment can back propagate to the first segment especially when the difficulty of the task is challenging.

**On-line adjustments.** More recently, Vindras and Viviani refuted the hypothesis that fast sequences of movements are planned as a whole before movement initiation. In their experiment, Viviani and Vindras (2005) asked participants to point to either one or two visual targets aligned along the mid-sagittal axis in a horizontal plane. They found differences in duration (one-target advantage), but velocity profiles and spatial accuracy did not depend on the characteristics of the second movement. Hence, the results were inconsistent with the assumption that two-step sequences are planned as a whole. Instead, findings were in keeping with an alternative hypothesis of an on-line strategy: The preparation of the second segment takes place during the execution of the first segment. It is plausible that both strategies of motor planning are possible within the cognitive brain and that they compete with each other in function of the constraints imposed by the task and/or the context. The more conservative on-line strategy would be selected when the motor control system does not have access to information relative to the initial conditions of each successive step. Thus, in the vast majority of everyday gestures the on-line strategy would be favoured because it is difficult to assume that the initial conditions such as timing, position, velocity and muscle tension information are defined and/or available a priori of movement initiation (Vindras & Viviani, 2005).

**Anticipatory Planning.** In a very recent research, it has been proposed nevertheless that *anticipatory modifications* will occur during the first movement when participants know the exact location of the second target of a two-step movement. In order to test this hypothesis, Herbort and Butz (2009) asked participants to execute a two-step hand movement sequence prior to a choice reaction signal. Results indicated that participants were able to initiate the

sequence quicker if they had advance information about the position of the second target of the sequence compared to the situation in which no advance information was provided. The authors concluded that information about the second target prior to action initiation was thus important for the preparation of the global motor plant even if the specification of the constraints of the first element were not yet known. Overall, the theory of *anticipatory planning of movement* claims that information about late targets in a sequence can be processed early and independently from that information known about the initial motor element, in order to define and optimize the general action plant layout (Oliver Herbert & Butz, 2009). Hence, sequential actions that are prepared in an anticipatory fashion should reveal back propagation of the constraints of later motor segments only if these constraints are important to incorporate early on during the motor planning phase, to insure that action goals are reached. In the following section, we consider this question for a different class of arm movements, i.e., those consisting in object manipulation.

## **Sequential object manipulation**

**Reaching for an object.** In everyday life, pointing and rapid arm movements are most commonly involved in actions that end with the manipulation of an object. Recently, a large number of authors have given a strong interest to object manipulation tasks because of its obvious ecological value in the better understanding of intentional object-directed actions. When we plan to interact with an object, we usually drive our hand in the direction of the object while at the same time we open our hand wide enough to grasp it. These two components (i.e., transport and manipulation) of the reaching movement have been well described in early studies by Jeannerod and collaborators (M. Jeannerod, 1984, 1988). More specifically, Jeannerod demonstrated that the temporal sequence and the coordination of the transport and the manipulation components are achieved by a centrally generated temporal pattern that remains uninfluenced by on-line visual feedback control. Indeed, during the typical bell-shaped velocity profile of the wrist, the fingers begin to shape. This process of pre-shaping first involves a progressive opening of the grip with straightening of the fingers,

followed by a closure of the grip until it matches object size. The point in time at which grip size is the largest is a clearly identifiable landmark that occurs well before the fingers come into contact with the object, e.g, at about 75% of total movement time (M. Jeannerod, Arbib, Rizzolatti, & Sakata, 1995). However, even though the coordination of reach and grasp has a strong temporal pattern, different effects of physical properties of the object on the reach-to-grasp initiation and execution have been reported.

**Effects of object size and distance.** When changing the object distance, empirical results showed very predictable effects on the transport component. Wrist movement directed towards far objects had a longer latencies, accelerated to peak velocity more quickly, attained a higher peak velocity, a higher maximum height and lasted longer than movements directed towards near objects. In addition, and more surprisingly, when objects were placed at farther distances, the maximum opening of the hand increased (Jakobson & Goodale, 1991; Marteniuk, MacKenzie, Jeannerod, Athenes, & Dugas, 1987; Zaal & Bootsma, 1993).

Numerous experimental studies have now demonstrated that the variation of peak aperture to which the hand is opened during approach increases in a linear function with respect to the size of the object to be grasped (Athenes, 1992; Jakobson & Goodale, 1991; Marteniuk, Leavitt, MacKenzie, & Athenes, 1990; Marteniuk et al., 1987). For example, Bootsma and colleagues (1994) showed that both height and width of the object affected peak hand aperture. Also, increasing object width reduced the spatial accuracy demands on the transport component, allowing a faster movement to emerge. As in rapid aiming movements, the effect of object width and movement amplitude also affects movement duration, influencing primarily the acceleration phase of the first reach segment. The effects of object width and movement amplitude were found to be combined with those effects predicted by Fitts' law, allowing a generalization of the latter to the transport component in prehensile actions (Bootsma, Marteniuk, MacKenzie, & Zaal, 1994; Zaal & Bootsma, 1993). Finally, effects of terminal accuracy constraints were shown to exist in one-segment manipulative actions. For example, Milner & Ijaz asked their participants to place a peg in various diameter holes. If the

diameter of the hole was large relative to the diameter of the peg, the tangential velocity profile of hand trajectories was symmetric and bell-shaped. As the diameter of the hole was reduced, peak velocity decreased, overall movement duration increased and the proportion of movement spent in deceleration increased systematically (Milner & Ijaz, 1990).

Taken together, the results reviewed here for one-segment object manipulative actions resemble those reported above for pointing movements: significant contextual effects on movement kinematics. In addition, a strong relationship between transport and grasp components are often reported. As compared to rapid aiming movements on which accuracy only affects movement times, in object manipulation movements, accuracy and speed constraints seem to affect simultaneously the global planning of the reach and of the grasp segments. This is what we will refer in the following as *coupled planning*.

**Reaching to use an object.** Sequential motor actions are dominant in most of our everyday movements, especially when interacting with another person. If you want to give an object to someone, first you must reach for the object, grasp it with a stable grip, transport the object across a certain distance before finally placing it safely in that person's hand. During the last decades, researchers have investigated the motor kinematics associated to sequential object manipulation and have found contextual effects on movement kinematics depending on both the context and the terminal motor goal of the action sequence.

In an early study, Marteniuk et al. (1987) asked participants to reach for an object and to either fit it into a similarly sized opening or to throw it away. The first reach-to-grasp phase revealed lower peak velocities and longer deceleration times for the „fit’ than the „throw’ condition (Marteniuk et al., 1987). When changing the distance to the second target, Gentilluci et al. (1997) reported an increase in peak velocity of the first segment. An increase in maximal finger aperture was also described when the second target was located in a far position compared to a near position. These results were interpreted within the hypothesis of the back-propagation of terminal constraints on the anticipatory planning of the global sequence, with motor planning taking account not only the constraints of the first segment but



also of all external constraints affecting the terminal motor goal. In other terms, the initial planning of the reach and grasp segments of the sequence were shown to be influenced by the visual analysis of the external proprieties of the immediate, intermediate and final targets of the motor sequence (Gentilucci, Negrotti, & Gangitano, 1997).

More recently, two studies have however set a doubt on the existence of the back-propagation effect for sequential pick and place actions through the assessment of the early effects of target size on movement kinematics (Johnson-Frey, McCarty, & Keen, 2004; Mason, 2007). In both cases, results suggested that despite substantial differences in the accuracy demands of the terminal object placement, no differences were found for the kinematics of the first segment. It was then suggested that kinematics of the first segment will only be affected with substantial changes to the second segment of the sequence, i.e., a change in the *nature* of the motor goal is required: pick to place; pick to lift; pick to toss (Ansuini et al., 2008; Armbrüster & Spijkers, 2006). It is to note however that in both experiments, the factors constraining the sequential actions were weak: participants were required simply to perform the task at preferred speed. Furthermore, terminal accuracy was not measured. Hence, the authors could not verify that the absence of differences in the kinematic patterns was not due to differences in the terminal accuracy performances of placing the object precisely upon the terminal pad, especially when contrasting small/large pad, and fast/slow speed constraints, respectively.

In the present series of studies, we present data confirming the back-propagation hypothesis in pick and place object manipulation sequences. We furthermore show that the strength of the constraints both in space and in time that is set upon the participants performing the task is crucial when wanting to reveal the use of the global anticipatory planning strategy. In a first experiment, we introduce a new methodology that takes advantage of real-time recordings for the study of movement kinematics. We then ask whether the constraints on the movement and the subjective feeling of interactivity are necessary factors to lead participants to adopt a global planning strategy for object manipulation. In a second

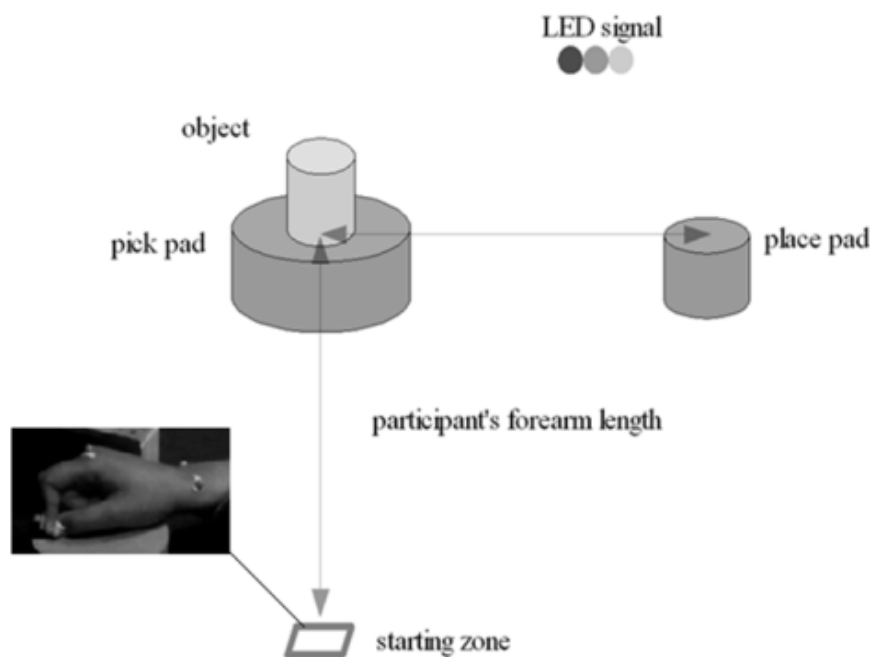
experiment, using the real-time setup, we question to what extent the spatial constraints for the place movement are *back-propagated* and may affect the planning of the first reach to grasp element. We were specifically interested here to confirm coupled planning in motor sequences by showing that by increasing the terminal accuracy demands, the two parts of the sequence are more tightly correlated. Finally, in a third experiment, we addressed the question of the limits of the use of the coupled planning strategy. We here investigated the effects of introducing an unpredictable time delay between the two elements of the sequence in order to assess the importance of predictability for anticipatory planning.

## **EXPERIMENT 1: Real time to monitor compliancy to speed constraints**

The question here was to assess whether the effects of speed and accuracy constraints on action planning may be different when the constraints are (1) imposed by the environment through a real-time interactive system and (2) simply through verbal instructions. Our hypothesis is that Real-Time will induce a specific cognitive state of anticipatory planning because of the high control set upon the participants' compliance to follow instructions. This state of compliancy should lead to higher levels of focused attention during the experimental session, a greater subjective feeling of interactivity and an increased coupling of movement kinematics. Following the results previously published in the pointing literature (Rand & Stelmach, 2000), these patterns of results should be observed especially during fast movements, because of the increase difficulty in the task but also because of the lack of time that prohibits the use of online sensory information.

## **Method**

**Participants.** Forty adults (mean age:  $26.14 \pm 6.06$  years, range = 21 - 45 years) participated in the study and were separated in two groups. Eighteen performed the task under the Verbal condition and twenty-two performed under the Real-Time condition. All subjects were right handed (Oldfield, 1971) and had no prior knowledge of the experimental protocol and objectives. Subjects provided informed consent before participating in the experimental session that lasted approximately 45 minutes.



***Figure 1.*** Illustration of the experimental setup.

The participants' task was to reach and grasp the green wooden object between thumb and index finger in order to move it from the pick pad to the place pad. Participants started each trial by pinching index and thumb together within the starting zone. Under Real-Time induction, three LEDs were placed close to the boundaries of the workplace in order to indicate the critical phases of each trial: red (“wait”); orange (“ready”); green (“go”). Under Verbal instructions, the LEDs were turned off.

**Apparatus and Software.** Before the start of the session, the length of each participant’s forearm was measured from the elbow to the tip of the index. This distance was taken as the reference distance to set the relative positions of pick and place pads (*Figure 1*).

Participants were seated comfortably facing a table in a dark and silent room. For each trial, participants started by placing the right hand close to body mid-line, within a starting zone that was delimited by tape placed directly on the tabletop. Pick and place pads were red wooden cylinders of 7.5cm high. The pick pad was of 15 cm of diameter; Place pads had variable diameters of 9 cm and 6 cm, depending on the condition, which in the following sections will be referred to as “large” and “small” place pads, respectively. The object that was to be manipulated by the participants was a green wooden cylinder of width 6 cm and height of 7.5 cm that was placed precisely in the centre of the pick pad, before the start of each trial. Kinematics data for the subjects' movements were recorded with 4 Oqus infra-red Cameras (Qualisys). Infrared reflective markers were placed on the index (base and tip), the thumb (tip), the wrist (scaphoid and pisiform) of each participant, as well as on the object (*Figure 1*). Cameras were calibrated before each recording session, allowing the system to reach a standard deviation smaller than 0.2 mm for any absolute position in the recording area, at a 200 Hz sampling frequency.

**Experimental Procedure.** The participants' task was to reach and grasp the green wooden object between thumb and index finger in order to move it from the pick pad to the place pad. Participant started each trial by pinching index and thumb together within the starting zone (*Figure 1*). After a random delay (1 to 3s), they initiated their pick and place sequence. Participants were required to perform the movements at two different speeds (fast vs. preferred) and with two different terminal accuracy constraints (large vs. small place pad) following a random block design. The instructions for fast speed were: “Perform at maximum possible speed”; no specific instructions were given for preferred speed. Four series of 5 successful trials were presented. After 5 trials a pause was systematically proposed to the participant. Depending of the group, participants performed the pick and place tasks under two different conditions.

**VERBAL INSTRUCTIONS.** Participants were required to initiate their movements on the verbal “Go” signal given by the experimenter. After 4 seconds, an auditory signal indicated

the end of the trial. After each trial, the experimenter placed the object back on the pick pad, then asked the subjects to place their hand inside the starting zone and to wait for the next “Go” signal.

**REAL-TIME INDUCTION.** Three LEDs were placed close to the boundaries of the workplace in order to indicate the critical phases of each trial: red (“wait”); orange (“ready”); green (“go”). The red light indicated the inter-trial delay of 4 seconds. As in the verbal instruction condition, with the yellow light, a random delay of 1 to 3 seconds was included to avoid any anticipatory effects for the start of the trial. When the light turned green, the subjects were to initiate their pick and place sequential actions. After each trial a sound was played, either a low pitch sound (“yoshi’s tongue”) for bad trials, or a high pitch sound (“acquired coin”) for successful trials. These sounds were selected from a video game (Super Mario Bros®).

Participants who performed the Real-Time condition were asked at the end of the session to fill in a questionnaire about their interactive feeling with the system. Two questions were asked: “Did you feel in interaction with the system?”; “Do you think that the reinforcing sounds at the end of each trial made the task more pleasant?”. The participants answered using 10 cm long analogical scales with far left coding for “No not at all” and far right coding for “Yes, absolutely”.

**Real-Time modeling and trial selection.** Before the start of the session, participants were asked to make five reach and grasp trials in order to constitute a real-time model of their hand. Data was transferred and analysed through the use of custom software (MATLAB). Real-time analysis of the data provided the means to detect whether all of the markers were fully recorded throughout the trial. If any marker was missing on any of the recorded frames, the trial was excluded from the analysis and repeated. In order to constrain the subjects actions in real-time, two parameters were verified. First, an analysis of the end position of the object was implemented. The object end position had to be placed within a virtual circle of radius of 5 mm or 15 mm, for the small and the large pad, respectively. If these constraints

were not respected, the trial was excluded from the analysis and repeated. Second, an analysis of the peak velocity was implemented. Under Preferred speed constraints, the amplitude of peak velocity of the first movement was required to be below  $1200 \text{ mm.s}^{-1}$ . Under Fast speed condition, the amplitude of peak velocity was required to be above  $1300 \text{ mm.s}^{-1}$ . If movements did not meet these requirements, the low tone was emitted and the trial was excluded and repeated. Under verbal instruction, the trial selection was manual. The experimenter decided whether the Pick action was performed “as fast as possible” and verified that the object was correctly stacked upon the place pad.

**Dependent variables and statistical analyses.** Raw data was filtered with a dual-pass 4th order Butterworth low-pass filter with a cut-off frequency of 15 Hz. A series of kinematic parameters were extracted from the recordings including Movement Times (MT) and Amplitude of Peak Velocities (APV). Mean values for the 15 trials in each condition and speed constraints were submitted to a repeated-measure ANOVA with Accuracy constraints (High; Low) and Condition (Real-Time; Verbal) as within factors. Two separate analyses were conducted on preferred speed and fast speed trials. An alpha level of 0.05 was set for all statistical analyses and post hoc analyses (Scheffé) were used when required.

## **Results and Discussion**

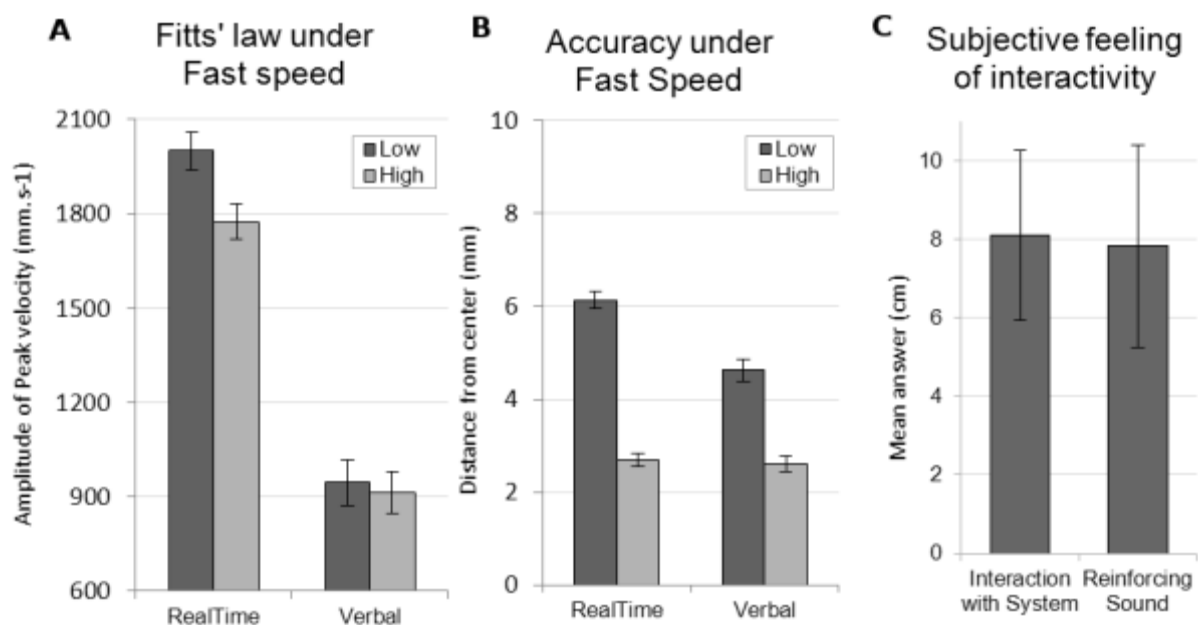
One of the most robust findings in the literature of motor control is the speed accuracy trade-off law, which states that if the target size is smaller, the movement towards this target will have a longer duration and a lower peak velocity (Fitts, 1954). However, in the case of a motor sequence, the question remains whether constraints that are set upon the second element back-propagate to impact the parameters of the first element of the sequence. The purpose of this first experiment was to examine whether participants would be more affected by the speed and accuracy constraints under a Real-Time induction condition as compared to a Verbal instruction condition. As such, the analyses focused on the kinematics of the first element only.

**Movement kinematics.** At Preferred speed, the kinematics of the first element was not affected by the accuracy constraints set upon the second element of the sequence: similar kinematic patterns were observed for movements towards small and large end-pads. Indeed, no effects of Accuracy were revealed on APV1,  $F(1,36) = 0.005$ ,  $p = .943$ , and on MT1,  $F(1,36) = 0.370$ ,  $p = .547$ . However, a general effect of condition was revealed: Higher peak velocities (APV1:  $M = 1002$ ,  $SD = 88$  mm.s<sup>-1</sup> vs.  $M = 759$ ,  $SD = 103$  mm.s<sup>-1</sup>;  $F(1,36) = 71.339$ ,  $p < .001$ ,  $\eta^2_p = .66$ ) and smaller movement durations (MT1:  $M = 897$ ,  $SD = 140$  mm.s<sup>-1</sup> vs.  $M = 784$ ,  $SD = 179$  mm.s<sup>-1</sup>;  $F(1,36) = 4.740$ ;  $p = .036$ ,  $\eta^2_p = .12$ ) were observed when performing under Real-Time than under Verbal condition. At Fast speed, movements were overall faster under low Accuracy constraints set upon the second element and under Real-time condition. For APV1, the main effect of Accuracy was significant,  $F(1,36) = 24.404$ ,  $p < 0.001$ ,  $\eta^2_p = .40$ , as well as the main effect of Condition,  $F(1,36) = 123.82$ ,  $p < 0.001$ ,  $\eta^2_p = .77$ . Similar patterns of results were observed for MT1 with a significant main Accuracy effect,  $F(1,36) = 3.797$ ,  $p = .059$ , and a tendency towards a main Condition effect,  $F(1,36) = 3.319$ ,  $p = .077$ . Most importantly, the interaction Accuracy  $\times$  Condition was significant for both APV1,  $F(1,36) = 13.836$ ,  $p < .001$ ,  $\eta^2_p = .28$ , and MT1,  $F(1,36) = 4.425$ ,  $p = .042$ ,  $\eta^2_p = .11$ : peak velocities and movement durations were affected by the accuracy constraints only when subjects were performing under Real-Time condition (*Figure 2A*). Scheffé post-hoc analyses confirmed that under Real-Time condition participants were characterized by smaller velocity peaks when reaching for an object to be placed on a smaller pad (high Accuracy,  $M = 1775$ ,  $SD = 289$  mm.s<sup>-1</sup>) as compared to a larger pad (low Accuracy,  $M = 2001$ ,  $SD = 341$  mm.s<sup>-1</sup>),  $p < 0.001$ . Similarly, longer movement durations were observed in high Accuracy trials ( $M = 457$ ,  $SD = 70$  ms) than in low Accuracy trials ( $M = 408$ ,  $SD = 66$  ms),  $p = .033$ . A totally different pattern of results was observed under Verbal condition with an absence of Accuracy effects, for both peak velocity (high Accuracy,  $M = 914$ ,  $SD = 219$  mm.s<sup>-1</sup>; low Accuracy,  $M = 945$ ,  $SD = 180$  mm.s<sup>-1</sup>),  $p = .886$ , and movement duration (high Accuracy,  $M = 479$ ,  $SD = 104$  ms; low Accuracy,  $M = 481$ ,  $SD = 117$  ms),  $p = .999$ .

**End position errors.** At Preferred speed, end position errors were affected by the accuracy constraints set upon the second element of the sequence,  $F(1,36) = 43.645$ ,  $p < .001$ ,  $\eta^2_p = .55$ . Participants showed larger errors when placing the object under low accuracy constraints ( $M = 3.01$ ,  $SD = 0.97$  mm) than under high accuracy constraints ( $M = 1.92$ ,  $SD = 0.70$  mm). More importantly, an effect of Condition was revealed,  $F(1,36) = 5.305$ ,  $p = .027$ ,  $\eta^2_p = .13$ , showing that participants made larger errors under Verbal instructions ( $M = 2.73$ ,  $SD = 0.59$  mm) than under Real-Time induction ( $M = 2.28$ ,  $SD = 0.50$  mm). A small trend for the Accuracy  $\times$  Condition interaction was observed,  $F(1,36) = 3.133$ ,  $p = .085$ ,  $\eta^2_p = .08$ . Scheffé post-hoc analysis confirmed that under low accuracy constraints, participants showed larger errors with Verbal instructions ( $M = 3.44$ ,  $SD = 1.08$  mm) than with Real-Time induction ( $M = 2.70$ ,  $SD = 0.77$  mm),  $p = .046$ ; while no significant differences were revealed under high accuracy constraints,  $p = .961$ . At Fast speed, a general effect of Accuracy was revealed,  $F(1,36) = 130.175$ ,  $p < .001$ ,  $\eta^2_p = .78$ , with larger errors under low accuracy constraints ( $M = 5.50$ ,  $SD = 1.54$  mm) than under high accuracy constraints ( $M = 2.65$ ,  $SD = 0.56$  mm). An Effect of Condition was also found,  $F(1,36) = 10.596$ ,  $p = .002$ ,  $\eta^2_p = .23$  but this time indicating that participants made smaller errors under Verbal instructions ( $M = 3.61$ ,  $SD = 0.75$  mm) than under Real-Time induction ( $M = 4.41$ ,  $SD = 0.75$  mm). This result indicates clearly that Real-Time induction revealed to be a powerful experimental setup to amplify the participants' compliance with task instructions. Finally, the Accuracy  $\times$  Condition interaction was significant,  $F(1,36) = 8.987$ ,  $p = .005$ ,  $\eta^2_p = .20$ . Post-Hoc Scheffé tests confirmed that under low accuracy constraints, participants showed smaller errors with Verbal instructions ( $M = 4.62$ ,  $SD = 1.73$  mm) than with Real-Time induction ( $M = 6.14$ ,  $SD = 1.02$  mm),  $p = .046$ ; while under high accuracy constraints, no significant differences were observed,  $p = .997$  (*Figure 2B*).



**Evaluating the degree of interactivity.** In order to better understand what could have occurred during Real-Time induction, we asked participants to note their overall feeling of interactivity with the Real-time system. On a 0-10 scale with 10 indicating high interactivity, our results showed that 71% of participants answered within the highest quarter ( $M = 8.10$ ,  $SD = 2.17$ ) indicating a high feeling of interactivity. Participants also found the reinforcing sounds pleasurable with 12 participants out of 14 rating the sounds above 7.5 ( $M = 7.82$ ,  $SD = 2.57$ ). Results are presented in Figure 2C.



*Figure 2.* Performance parameters for fast movements and subjects feeling of interactivity: Real-Time vs Verbal Instruction (*Experiment 1*).

Amplitude of Peak Velocity of the reaching movement (A) and End position Error (B) at Fast Speed recorded under Real-Time and Verbal-Instruction conditions. Light-grey bars indicate means (standard error) for high accuracy constraints (small place pad). Dark-grey bars indicate means (standard error) for low accuracy constraints (large place pad). Mean answers (standard error) reported on analogical scales for the subjective feeling of interactivity (C).

**Conclusions.** The results for this first experiment confirmed previous reports showing that at preferred speed, the kinematics of the first element is not affected by the accuracy

constraints imposed to the second element of a 2-element motor sequence (Johnson-Frey et al., 2004; Mason, 2007). This was true for both verbal and real-time conditions, even if participants tended to move faster under real-time induction. When planning to perform the task as fast as possible, the *back propagation* of the speed-accuracy constraint was revealed in the first element of the sequence but in the real-time condition only. These findings reproduce those previously reported in the pointing literature (Rand et al., 1997; Rand & Stelmach, 2000). In classic object manipulation experiments, actions are usually performed at preferred speed. Hence, in the present case, we imposed the same speeded condition than in pointing experiment and our results clearly suggest that the well-known speed-accuracy trade-off can emerge for manipulative tasks but only when *accuracy* and *speed* constraints are close to the maximum of what a participant can perform. Furthermore, Real-Time provided participants with reliable and immediate feedback about their performance. This interactive setup revealed more stable results with lower spatial errors and faster movements. Hence, real-time may be a valuable setup to use to enhance the effects of experimental conditions.

Taken together, results reported in this first experiment confirmed that planning sequential pointing on the one hand, and the planning of sequential manipulative actions on the other hand are subjected to similar production laws of action (Shadmehr & Wise, 2005). The discrepancies seen in the literature of sequential movements between pointing and reaching may simply be due to the use of different experimental paradigms characterized by contrasting control parameters. As the objective of our contribution was to analyse the coupling properties of subsequent movements during pick to place movements, it was necessary for us to develop a protocol that would induce *true* speed-accuracy constraints that could be manipulated reliably. The results of experiment 1 demonstrate that the real-time induction developed using the Qualysis® cameras and Matlab® software provides the means to propose a highly interactive system that offers furthermore a pleasurable experience to our participants. Thanks to the real-time procedure ( $< 0.25$  mm;  $< 5$  ms), it was in addition possible to monitor the accuracy of the participant's performances in order to modulate

constraining situations when required. For these various reasons, the Real-Time induction only was used in experiments 2 and 3.

## **EXPERIMENT 2: Kinematic analysis of the pick-and-place elements.**

The question here was to assess the strength of the coupling between two subsequent elements of a motor sequence, i.e., how the different kinematic parameters of the pick-and-place segments are correlated. More specifically, we will question the contextual conditions needed to induce the coupled mode of action planning. Maintaining the Real-time condition only, we introduced an intermediate accuracy constraint in order to test whether the interaction effects observed between *speed* and *accuracy* were a continuous or a more discrete phenomenon. In the present case, we hypothesized that results would follow a more continuous modulation with a log linear relation between speed and accuracy, as suggested by the Fitts' law in cyclical movements (Fitts, 1954; Guiard, 1993, 1997).

### **Method**

**Participants.** Twenty-two adults (Age:  $M = 20.24$ ,  $SD = 2.57$  years, range = 17-28 years) participated in the study and were right handed (Oldfield, 1971). These subjects had no prior knowledge of the experimental protocol (did not participant in study 1) and provided informed consent before participating in the experimental session that lasted approximately 45 minutes.

**Apparatus & Software.** The apparatus was similar to that presented in experiment 1 with an additional place pad of 7.5cm diameter. Thus in this experiment, the place pad could be of diameters 9 cm, 7.5 cm or 6 cm, depending on the accuracy constraints. This will be referred to as “low”, “mid” and “high” in the following sections, respectively.

**Procedure.** The participants' task was to reach and grasp an object between thumb and index finger in order to move it from the pick pad to the place pad. Participants started

each trial by pinching index and thumb together within the starting zone. The LEDs indicated the critical phases of each trial: red (“wait”); orange (“ready”); green (“go”). For the orange light, a random delay of 1 to 3 seconds was included to avoid anticipatory effects. When the light turned green, the subjects were to initiate the pick and place sequence actions either at preferred or at fast speed, depending on the speed condition.

**Real-Time Modeling.** The procedure used in this second experiment was similar to that presented in experiment 1. The virtual circles were of radius of 5 mm, 7.5 mm and 15 mm. Real-time analysis of the reach movement (1<sup>st</sup> element) provided the means to verify the subjects’ speed. For preferred speed, the amplitude of peak velocity was required to be below 1200 mm.s<sup>-1</sup>. For fast speed, the amplitude of peak velocity was required to be above 1300 mm.s<sup>-1</sup>. If these constraints were not respected, the trial was excluded from the analysis and repeated at the end of the block.

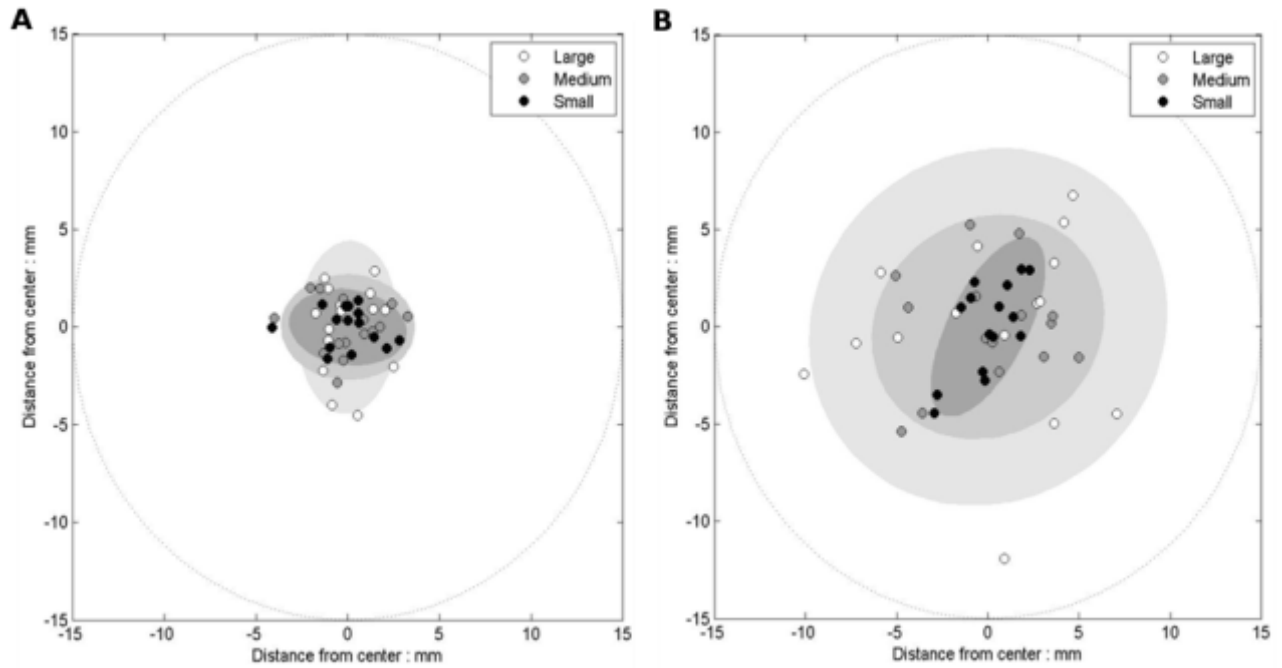
**Dependent kinematic variables and statistical analyzes.** Raw data was filtered with a dual-pass fourth order Butterworth low-pass filter with a cutoff frequency of 15 Hz. For the grasping action, the Amplitude of maximum Grip Aperture (AGA), Time of maximum Grip Aperture (TGA) and amplitude of peak Velocity of Grip Aperture (VGA) were extracted from the recordings. For the reaching and placing actions, another series of parameters were extracted: Reaction Time (RT), Movements Time (MT) and Amplitude of Peak Velocity (APV), End position errors (Error), Time to Peak Velocity (TPV) and Percentage of Deceleration Time (%DT). Mean values for the 15 trials were submitted to a repeated-measure ANOVA with Accuracy constraints (High; Mid; Low) and Speed (Preferred; Fast) as within factors. An alpha level of 0.05 was set for all statistical analyses and post hoc analyses (Scheffé) were used when required.

## **Results and Discussion**

The purpose of experiment 2 was to investigate whether movement kinematics of element 1 (pick) and element 2 (place) were correlated when planned and executed within the same

motor sequence. This would be suggestive of coupled planning for the execution of a global motor sequence. In order to infer about possible accuracy effects on movement coupling, we first tested whether the accuracy of the end-point object position was truly reflecting the constraints that were imposed. Second, we analysed the effects of speed and accuracy constraints of the 2<sup>nd</sup> element on the different kinematic parameters of the 1<sup>st</sup> element both for the grasp and the reach components. In a third and final section, we conducted a series of correlation analyses between the kinematic parameters of elements 1 (pick) and 2 (place).

**Typical patterns of endpoint distributions.** For the measurements of end-point distribution, ANOVA on Error revealed a main effect of Speed,  $F(1,21) = 179.18$ ,  $p < 0.001$ ,  $\eta^2_p = .90$ , indicating that errors were larger when subjects moved as fast as possible than at preferred speed. As expected, the main effect of Accuracy was also significant,  $F(2,42) = 115.47$ ,  $p < 0.001$ ,  $\eta^2_p = .85$ , showing that errors were smaller when subjects placed the object on the smallest pad ( $M = 2.3$ ,  $SD = 0.3$  mm) than when they placed the object on either the mid one ( $M = 3.3$ ,  $SD = 0.5$  mm) or the largest one ( $M = 4.4$ ,  $SD = 0.7$  mm). Nevertheless, the interaction was also significant,  $F(2,42) = 21.226$ ,  $p < .001$ ,  $\eta^2_p = .74$ , indicating that the Accuracy effect was stronger at Fast speed than at Preferred speed. Post-hoc Scheffé tests, on spatial *Error*, confirmed strong contrasts at Fast speed between High ( $M = 2.7$ ,  $SD = 0.4$  mm), Mid ( $M = 3.9$ ,  $SD = 0.6$  mm) and Low accuracy constraints ( $M = 6.1$ ,  $SD = 1.0$  mm). However, spatial errors were similar at Preferred speed for High ( $M = 1.87$ ,  $SD = 0.5$  mm), Mid ( $M = 2.6$ ,  $SD = 0.8$  mm) and Low accuracy constraints ( $M = 2.7$ ,  $SD = 0.8$  mm) suggesting that at preferred speed, subjects were not taking into account that larger accurate placing movements could be made with larger place pads. Results from a typical subject (individual TQ) are presented in Figure 3.



*Figure 3.* Terminal position errors – Typical Individual (*Experiment 2*).

End-point distributions observed in a typical subject (TQ) are presented under Preferred speed (A) and Fast speed (B) conditions. White, grey and black filled circles illustrate the different end-positions recorded during Low, Mid and High accuracy conditions, respectively. The grey filled areas around the circles correspond to the 95% confidence ellipses obtained for each condition separately.

**Kinematic measures of reach.** As expected, results revealed a significant main effect of Speed on three key kinematic parameters. First the maximum APV was higher at Fast Speed ( $M = 1878$ ,  $SD = 302$  mm) than at Preferred speed ( $M = 995$ ,  $SD = 71$  mm.s<sup>-1</sup>),  $F(1,21) = 172.737$ ,  $p < .001$ ,  $\eta^2_p = .89$ . Second, the MT was shorter at Fast Speed ( $M = 428$ ,  $SD = 60$  ms) than at Preferred speed ( $M = 875$ ,  $SD = 137$  ms),  $F(1,21) = 305.041$ ,  $p < .001$ ,  $\eta^2_p = .93$ . Finally, the maximum peak acceleration (APA) was higher at Fast Speed ( $M = 13134$ ,  $SD = 2934$  mm.s<sup>-2</sup>), than at Preferred speed ( $M = 4659$ ,  $SD = 882$  mm.s<sup>-2</sup>),  $F(1,21) = 199.068$ ,  $p < .001$ ,  $\eta^2_p = .90$ . More interestingly, a main effect of Accuracy on the second element of the sequence was observed on the first element for APV<sub>1</sub>,  $F(2,42) = 30.796$ ,  $p < .001$ ,  $\eta^2_p = .59$ , showing that participants peak velocity was higher when reaching to place an object on a large pad ( $M = 1502$ ,  $SD = 178$  mm.s<sup>-1</sup>) compared to that measured for both the medium pad ( $M = 1422$ ,  $SD = 159$  mm.s<sup>-1</sup>) and the small pad ( $M = 1380$ ,  $SD = 143$  mm.s<sup>-1</sup>).

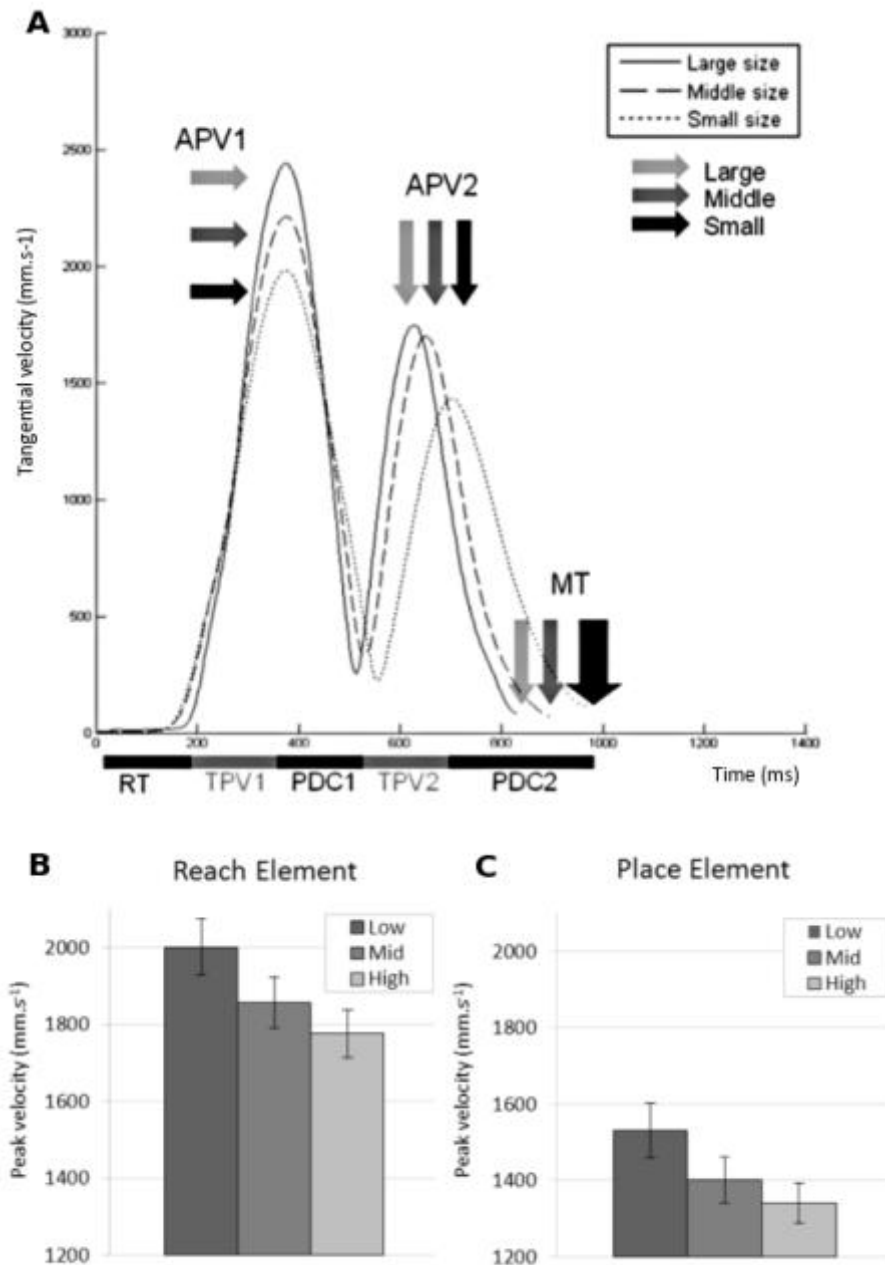
Similar back propagation effects were obtained for MT<sub>1</sub>,  $F(2,42) = 5.438$ ,  $p = .008$ ,  $\eta^2_p = .21$ , and for APA<sub>1</sub>,  $F(2,42) = 8.084$ ,  $p = .001$ ,  $\eta^2_p = .28$ . The details of the means and standard deviation of APV<sub>1</sub>, MT<sub>1</sub> and APA<sub>1</sub> are reported in Table 1. The interaction between Speed  $\times$  Accuracy interaction was significant on all three parameters of this first pick element (APV<sub>1</sub>:  $F(2,42) = 29.266$ ,  $p < .001$ ,  $\eta^2_p = .58$ ; MT<sub>1</sub>:  $F(2,42) = 3.665$ ,  $p = .034$ ,  $\eta^2_p = .15$ ; APA<sub>1</sub>:  $F(2,42) = 5.184$ ,  $p = .010$ ,  $\eta^2_p = .20$ ). More specifically, post-hoc Scheffé revealed significant differences between the Accuracy constraints at Fast speed. For example, mean APV<sub>1</sub> was higher when reaching for an object that was to be placed on the large pad ( $M = 1999$ ,  $SD = 337$  mm.s<sup>-1</sup>) compared to the medium pad ( $M = 1856$ ,  $SD = 309$  mm.s<sup>-1</sup>),  $p < .001$ , which was itself higher than the movements made towards objects that was to be placed on the small pad ( $M = 1767$ ,  $SD = 282$  mm.s<sup>-1</sup>),  $p = .013$ . No such effects were found at Preferred speed, with an absence of differences for movement kinematics of reaches towards an object to be placed on the large ( $M = 1002$ ,  $SD = 88$  mm.s<sup>-1</sup>), medium ( $M = 988$ ,  $SD = 84$  mm.s<sup>-1</sup>), and small place pad ( $M = 991$ ,  $SD = 86$  mm.s<sup>-1</sup>),  $p > .99$ . Similar patterns of results were obtained for both MT<sub>1</sub> and APA<sub>1</sub>.

		Preferred			Fast			Anova		
		Low	Mid	High	Low	Mid	High	p Speed	p Accuracy	p Speed*Acc.
								( $\eta^2_p$ )	( $\eta^2_p$ )	( $\eta^2_p$ )
Reaction	Time	291 (68)	310 (104)	284 (63)	199 (28)	217 (51)	222 (46)	<0.001 (0.89)	0.483	0.510
(ms)										
Movement	Time	1951 (321)	2033 (407)	2048 (335)	805 (135)	909 (125)	986 (167)	<0.001	0.069	0.707
(ms)										
End Position		2.7 (0.2)	2.6 (0.2)	1.9 (0.1)	6.1 (0.2)	3.9 (0.1)	2.7 (0.1)	<.001 (0.90)	<.001 (0.85)	<.001 (0.74)
Error (mm)										
Grasp Formation										
MGA (mm)		119 (1)	119 (2)	119 (2)	147 (3)	143 (3)	141 (3)	<.001 (0.82)	.036 (0.15)	<.001 (0.29)
TGA (ms)		586 (18)	588 (20)	592 (19)	270 (11)	289 (11)	30 (12)	<.001 (0.92)	.043 (0.14)	.169 (0.08)
VGA (mm.sec <sup>-1</sup> )		369 (25)	383 (25)	413 (28)	900 (48)	883 (47)	827 (44)	<.001 (0.84)	.781 (0.01)	.026 (0.16)
Reaching Movement										
APV <sub>1</sub> (mm.sec <sup>-1</sup> )		1002 (19)	988 (18)	992 (18)	2001 (73)	1858 (66)	1775 (62)	<.001 (0.89)	<.001 (0.59)	0.001 (0.58)
TPV <sub>1</sub> (ms)		385 (14)	395 (17)	384 (13)	214 (7)	226 (8)	235 (8)	<.001 (0.88)	.124 (0.09)	.175 (0.08)
MT <sub>1</sub> (ms)		897 (30)	925 (39)	903 (31)	408 (14)	441 (14)	457 (15)	<.001 (0.92)	.015 (0.18)	.105 (0.10)
%DT <sub>1</sub> (%)		56.8 (1.1)	57.2 (1.0)	57.3 (0.9)	47.6 (0.8)	49.1 (1.0)	48.8 (0.8)	<.001 (0.74)	.106 (0.10)	.461 (0.04)
Object Transport										
APV <sub>2</sub> (mm.sec <sup>-1</sup> )		686 (22)	680 (22)	701 (21)	1532 (71)	1402 (61)	1341 (53)	<.001 (0.88)	<.001 (0.42)	<.001 (0.57)
TPV <sub>2</sub> (ms)		663 (23)	687 (32)	645 (20)	354 (9)	385 (12)	393 (11)	<.001 (0.90)	.048 (0.13)	.008 (0.20)
MT <sub>2</sub> (ms)		1396 (49)	1444 (63)	1425 (48)	601 (18)	689 (22)	751 (22)	<.001 (0.92)	<.001 (0.32)	.006 (0.22)
%DT <sub>2</sub> (%)		52.3 (0.9)	52.5 (0.8)	54.6 (0.7)	41.2 (0.9)	43.8 (0.7)	47.0 (1.3)	<.001 (0.85)	<.001 (0.41)	.007 (0.21)

***Table 1.*** Kinematic Parameters and mean group results (*Experiment 2*).

Main experimental effects obtained for the kinematic measurements of Experiment 2. Means (standard errors) are shown for the results obtained in Fast and Preferred speed conditions, under low, mid or high accuracy constraints. Abbreviations: MT = Movement Time, RT = Reaction Time, APV = Amplitude of Peak Velocity, TPV = Time to Peak Velocity, %DT = percentage of Deceleration Time, MGA = Maximum Grip Aperture, TGA = Time to maximum Grip Aperture, VGA = amplitude of peak Velocity of Grip Aperture.





**Figure 4.** Typical individual velocity profiles and Group results (*Experiment 2*).

A. Typical Fast Speed velocity profile obtained in a typical subject (JG). Full, dashed and dotted lines correspond to profiles for Low, Mid and High Accuracy conditions, respectively. Profiles were constructed with the means of the 15 trials obtained in each experimental condition. Note that the first peak of velocity is higher in Low condition (large pad) than in Mid (medium pad) and High (small pad) condition. B, C. Group results, means and standard errors are computed across all subjects. The Speed  $\times$  Accuracy interaction was significant for APV<sub>1</sub>:  $F(2,42) = 28.792$ ,  $p < .001$ ,  $\eta^2_p = .58$  and for APV<sub>2</sub>:  $F(2,42) = 27.761$ ,  $p < .001$ ,  $\eta^2_p = .88$ . Means and standard errors are presented in Table 1.

**Movement kinematics of grasping.** Results revealed a significant main effect of Speed on three key kinematic parameters. First, the maximum AGA was wider at Fast speed ( $M = 143.4$ ,  $SD = 13.9$  mm) than at Preferred speed ( $M = 119.0$ ,  $SD = 7.4$ mm),  $F(1,21) = 97.937$ ,  $p < .001$ ,  $\eta^2_p = .82$ . Second, mean TGA was earlier at Fast speed ( $M = 288$ ,  $SD = 49$ ms) than at Preferred speed ( $M = 588$ ,  $SD = 82$  ms),  $F(1,21) = 269.113$ ;  $p < .001$ ,  $\eta^2_p = .93$ . Finally, mean VGA was faster at Fast speed ( $M = 865$ ,  $SD = 185$  ms) than at Preferred speed ( $M = 388$ ,  $SD = 111$  mm.s<sup>-1</sup>),  $F(1,21) = 112.839$ ,  $p < .001$ ,  $\eta^2_p = .84$ . A global effect of Accuracy constraint of the second element was also observed on AGA,  $F(2,42) = 4.085$ ,  $p = .024$ ,  $\eta^2_p = .16$ , and TGA,  $F(2,42) = 4.899$ ,  $p = .012$ ,  $\eta^2_p = .19$ , but the span of the effects were small, i.e., within intervals of 3 mm and 20 ms, respectively, which could not be significantly confirmed with post hoc Scheffé measures.

**Correlation analysis between pick and place.** As a first descriptive approach, we conducted a principal components factorial analysis, using the 13 kinematic variables of the reaches for element 1 (pick) and for element 2 (place). Results revealed a strong correlation between MT<sub>1</sub> and MT<sub>2</sub> variances, on the one hand, and APV<sub>1</sub> and APV<sub>2</sub> variances, on the other hand. Both RT and Error, as well as other kinematic variances (MGA, TGA, VGA, TPV<sub>1</sub>, TPV<sub>2</sub>, %DT<sub>1</sub>, %DT<sub>2</sub>,) were independent of the distributions of the aforementioned kinematic parameters (Figure 5). Overall, these findings suggest coupled planning of the kinematics of sequencing reaching movements without modification of the general motor laws that guide reaching and grasping components. In the final section, we report the effects of condition on the motor elements that are encapsulated within a unique and global motor plant.

**APV<sub>1</sub>-APV<sub>2</sub>.** Results showed a significant global effect of Speed on the coefficient of determination ( $R^2$ ) of the linear regression between APV<sub>1</sub> and APV<sub>2</sub>. The  $R^2$  was stronger at Fast speed ( $M = 0.241$ ,  $SD = 0.139$ ) than at Preferred speed ( $M = 0.148$ ,  $SD = 0.105$ ),  $F(1, 21) = 5.510$ ,  $p = .029$ ,  $\eta^2_p = .21$ . No effects were found for Accuracy constraints,  $F(2, 42) = 0.075$ ,  $p = .928$ . The interaction between Speed and Accuracy was significant,  $F(2, 42) = 4.9108$ ,  $p =$

.012. However, the size of the effect was very small ( $<.05$ ) and was not confirmed with Post-hoc Scheffé tests.

**MT<sub>1</sub>-MT<sub>2</sub>.** Results showed a significant global effect of Speed on the coefficient of determination ( $R^2$ ) of the linear regression between MT<sub>1</sub> and MT<sub>2</sub>. The  $R^2$  was stronger for sequences performed at Fast speed ( $M = 0.258$ ,  $SD = 0.154$ ) than at Preferred speed ( $M = 0.148$ ,  $SD = 0.105$ ),  $F(1, 21) = 10.409$ ,  $p = .004$ ,  $\eta^2_p = .33$ . No effects were found for Accuracy,  $F(2, 42) = 0.451$ ,  $p = .640$ , and the interaction between Speed and Accuracy was not significant,  $F(2, 42) = 0.185$ ,  $p = .832$ . Figure 5 illustrates patterns of results obtained in a typical subject (subject BQ).

**Conclusions** The results for this second experiment demonstrated that when manipulating the constraints of the second element of a sequence (speed and/or accuracy), the kinematics of the first element are modified. This was observed even when the reaching movement was identical in all aspects: same distance, same object width, and same instruction. These findings suggest that the conditions that constrain the end point of a sequence are taken into account very early during motor planning and as such, affect the motor planning of the initial reaching movement. This is suggestive of a back-propagation phenomenon with the terminal constraints inducing modifications of the kinematics of initial movement of the sequence, with reduced velocities and acceleration phases in order to produce a terminal movement that will minimize endpoint position error. Our findings reproduce those previously reported in the reaching literature considering endpoint comfort, but are to our knowledge the first to confirm a coupled planning for Fitts' Law in a sequenced pick and place task. In the third and last experiment, we wanted to confirm the presence of a coupled planning through a protocol that would induce de-coupling between sequential movements by introducing an unpredictable timing factor within the motor sequence.

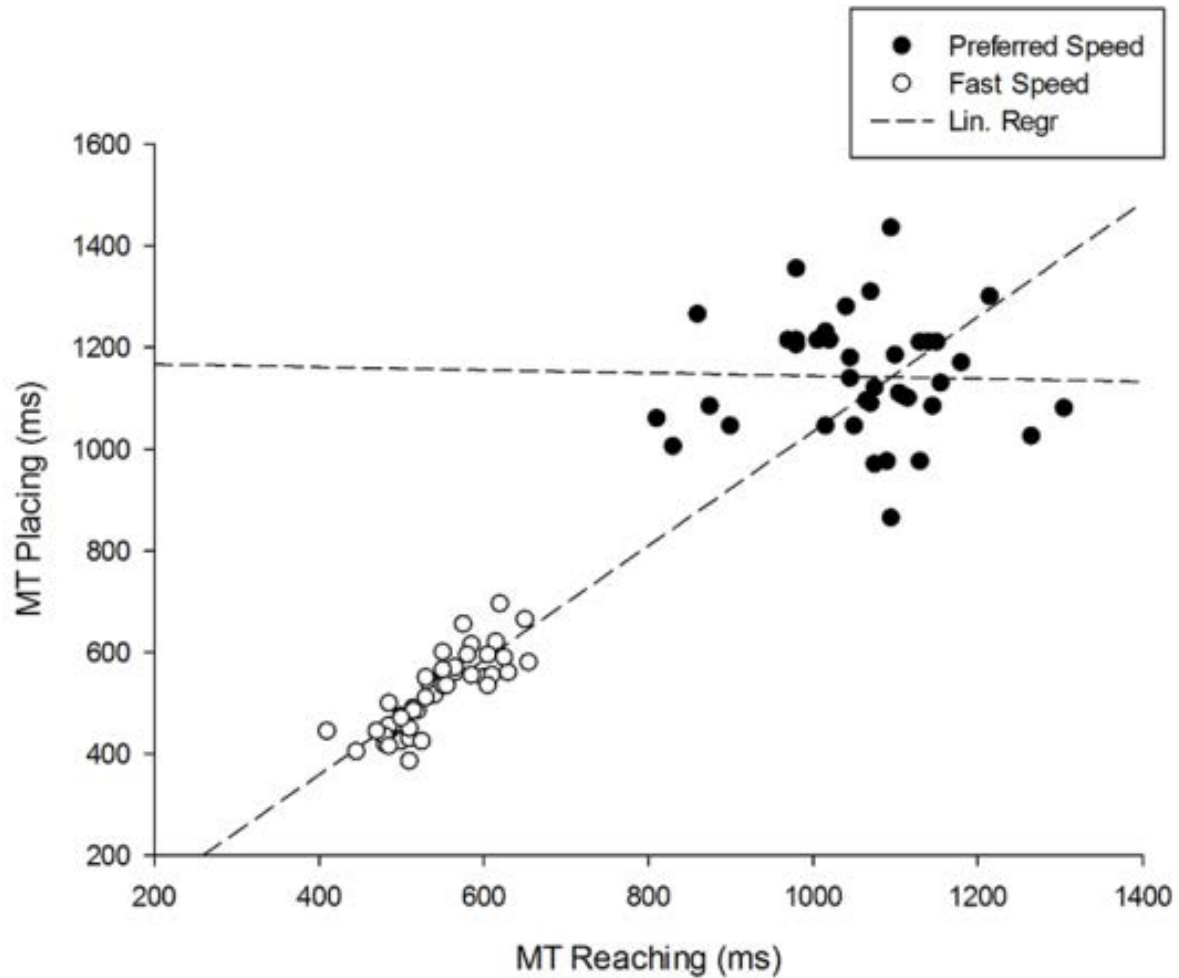


Figure 5. Scatterplots of  $MT_1$  and  $MT_2$  correlations under high and low accuracy constraints (*Experiment 2*).

Linear regressions of movement times (MT) are presented for placing movements performed at fast speed. White and black filled circles are individual trials for small end pad (high accuracy) and large end pad (low accuracy), respectively. Dashed lines are simple linear regressions lines (least squares fit). Note that in Preferred Speed condition  $MT_1$  and  $MT_2$  are poorly correlated ( $r^2 = 0.007$ ,  $p = 0.77$ ), while in Fast speed condition  $MT_1$  and  $MT_2$  are highly correlated ( $r^2 = 0.53$ ,  $p = .002$ ). Group results confirmed significant difference between coefficients of determination at Preferred and Fast speed (Preferred:  $r^2 = 0.12 \pm 0.03$  vs. Fast:  $r^2 = 0.34 \pm 0.06$ ,  $p = .007$ )

### EXPERIMENT 3: Temporal predictability required for coupled planning

The question here was to investigate the effect of introducing an unpredictable pause between two subsequent elements of a motor sequence. Since early experiments (Bernstein, 1967) to

the more recent observations, the concept of „co-articulation’ has been suggested to define the composition of motor elements bound to each other through time to achieve a goal through a sequential action. During the past twenty years, it has been reported that in well-trained motor sequences, the motor elements are influenced by anticipated adjacent elements, e.g., in piano playing (Engel, Flanders, & Soechting, 1997), handwriting (P. Viviani & Terzuolo, 1983), speech articulation (Macneilage, 1980) and typewriting (Soechting & Flanders, 1992). Most of these experiments have focused on the learning of the „co-articulation’ ability (Sosnik, Hauptmann, Karni, & Flash, 2004) or the „chunking’ strategy (Sakai, Kitaguchi, & Hikosaka, 2003) but very few have addressed the issues related to the predictability requirements needed to experience such motor learning. In the case of „space predictability’, studies have reported that strict accuracy imposed on an intermediate target will break the motor plant and lead to piecewise execution (Hesse & Deubel, 2010; Sosnik, Flash, Hauptmann, & Karni, 2007). The key question of „temporal predictability’ for coupled planning was recently addressed by Ansuini and collaborators (2009). Their results showed that when motor timing is prevented, the action sequence is no longer planned based on the end goal, but executed in discrete and independent action steps (Ansuini, Grigis, Adamovich, & Castiello, 2009), leading to a loss of motor fluency. However, in that study, movements were executed only at preferred speed and no correlation analyses between motor elements were proposed, probably because the *nature* of the first (reach-to-grasp) and the second element (pour into container) were different. In our third experiment, we will question the temporal conditions required for coupled planning and we will conduct a similar correlation analysis as that presented in Experiment 2 with an additional experimental factor: with or without time breaks.

## **Method**

Participants, apparatus and software. 20 subjects (Age:  $M = 23.64$ ,  $SD = 4.31$ , range = 18-29 years) participated in the study and were right handed (Oldfield, 1971). These subjects had no prior knowledge of the experimental setup (did not participate in experiments 1 and 2) and provided informed consent before participating in the experimental session that lasted

approximately 30 minutes. The apparatus and software were the same as that described in experiments 1 and 2, with the use of the large place pad only (diameter 9 cm).

**Procedure.** Two different conditions were proposed to the participants. During the “Pick and Place” condition, the participants had to reach and grasp the object between thumb and index finger in order to move it from the pick pad to the place pad. During the “Pick, Wait and Place” condition, the participants had to reach and grasp the object, then wait for the “go” signal. When the “go” signal was sent, participants were instructed to move the object from the pick pad to the place pad. As in experiment 2, three LEDs were placed close to the boundaries of the workplace in order to indicate the critical phases of each trial: red (“wait”); orange (“ready”); green (“go”). During the “Pick, Wait and Place”, when participants grasped the object, the lights turn back to orange during a random delay of 1 to 2 seconds. When the lights turned green again, the participants were instructed to move the object from the pick pad to the place pad. Subjects performed these conditions either at preferred or at fast speed, depending on the speed condition. Eight series of 10 successful trials were recorded. After 5 trials a pause was systematically proposed to the participants.

**Real-Time Modeling.** In order to verify that subjects respected each experimental constraints, a real-time analysis of the end position of the object was implemented. The object end position was required to be placed within a virtual circle of 15 mm. As in experiments 1 and 2, an initial set of recordings was used to calibrate the participants’ model hand position at initial position. In addition, for this experiment a model was also made of the participants’ hand position on the object. At this instant, the index finger was required to be within a virtual circle of 5 mm radius with respect to the calibrated object position. The participants’ index finger was required to remain within this virtual circle during the total time of the orange “wait” signal. If these constraints were not respected, the trial was excluded from the analysis and repeated at the end of the block. This real-time analysis provided the means to verify placing velocity both in the noPause and in the Pause trials.

Under the preferred speed conditions, no requirements were specified. Under the fast speed condition, the amplitude of peak velocity was required to be above 1000 mm.s<sup>-1</sup> for the element 2 (place); no requirements were placed on the first element of the sequence. Thus, accuracy, speed and temporal constraints of the second element of the sequence needed to be taken into account by the participants in order to perform successfully.

**Dependent variables and statistical analysis.** Raw data was filtered with a dual-pass fourth order Butterworth low-pass filter with a cutoff frequency of 15 Hz. For the reaching and placing actions, a series of motor parameters were extracted: Reaction Time (RT), Movements Time (MT) and Amplitude of Peak Velocity (APV), End position errors (Error). Mean values of the 15 trials were submitted to a repeated-measure ANOVA with Condition (noPause; Pause) and Speed (Preferred; Fast) as within factors. An alpha level of 0.05 was set for all statistical analyses and post hoc analyses (Scheffé) were used when required.

## Results and Discussion

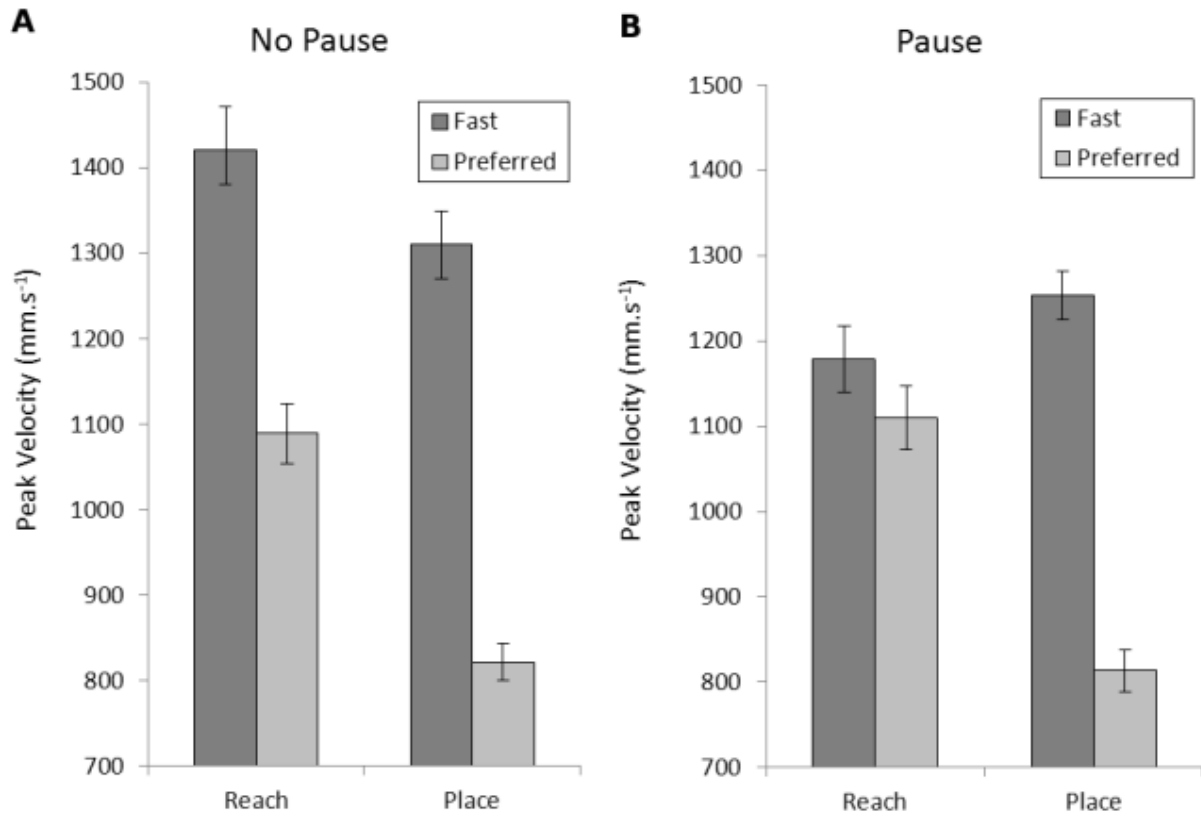
Typical patterns of endpoint distributions. For the measurements of end-point distribution, ANOVA on Errors revealed a main effect of Speed,  $F(1,18) = 56.995$ ,  $p < .001$ ,  $\eta^2_p = .76$ , indicating that errors were larger when subjects were instructed to perform the second element of the sequence as fast as possible ( $M = 5.6$ ,  $SD = 1.1$  mm) than at preferred speed ( $M = 3.6$ ,  $SD = 1.0$  mm). No interactions were found between Condition and Speed effects indicating that the differences between noPause and Pause were negligible ( $< 0.5$ mm) both at Preferred speed,  $p = .116$ , and at Fast Speed,  $p = .801$ . This later result suggests that when interrupting the movement, the final accuracy of sequence execution is not impacted.

**Kinematic measures of reach element.** Results revealed a significant main effect of Speed on two key kinematic parameters. Maximum APV<sub>1</sub> was higher,  $F(1,18) = 49.211$ ,  $p < .001$ ,  $\eta^2_p = .74$ , and MT<sub>1</sub> was shorter,  $F(1,18) = 61.980$ ,  $p < .001$ ,  $\eta^2_p = .77$  at Fast speed than

at Preferred speed. A global effect of Condition was also observed both on  $APV_1$ ,  $F(1,18) = 14.878$ ,  $p = .001$ ,  $\eta^2_p = .47$ , and  $MT_1$ ,  $F(1,18) = 42.555$ ,  $p < .001$ ,  $\eta^2_p = .70$ , indicating that when subjects reached for an object to be placed without interruption, the peak velocity was higher and the reaching movement duration was shorter. More importantly, the Condition  $\times$  Speed interaction was significant for both  $APV_1$ ,  $F(1,18) = 32.273$ ,  $p < .001$ ,  $\eta^2_p = .65$  and  $MT_1$ ,  $F(1,18) = 52.035$ ,  $p < .001$ ,  $\eta^2_p = .74$ . Post-hoc Scheffé confirmed a significant difference between Preferred and Fast speed under noPause condition, for  $APV_1$  ( $M = 1089$ ,  $SD = 146\text{mm.s}^{-1}$  vs  $M = 1420$ ,  $SD = 223\text{mm.s}^{-1}$ ),  $p < .001$ , and for  $MT_1$  ( $M = 667$ ,  $SD = 97\text{ms}$  vs  $M = 487$ ,  $SD = 100\text{ms}$ ),  $p < .001$ . Under Pause condition, post-hoc Scheffé revealed an absence of differences between preferred and fast speed both for  $APV_1$  ( $M = 1110$ ,  $SD = 161\text{mm.s}^{-1}$  vs  $M = 1168$ ,  $SD = 170\text{mm.s}^{-1}$ ),  $p = .40$  and for  $MT_1$  ( $M = 712$ ,  $SD = 82\text{ms}$  vs  $M = 703$ ,  $SD = 105\text{ms}$ ),  $p = .97$ . This suggests that when the sequence is interrupted by a time break, the reach for element is not affected by the speed constraints set upon the second element (place) of the sequence.

**Kinematic measures of place element.** For the Place element, results revealed a significant main effect of Speed on two key kinematic parameters. Maximum  $APV_2$  was higher at Fast speed ( $M = 1305$ ,  $SD = 131\text{mm.s}^{-1}$ ) than at Preferred speed ( $M = 824$ ,  $SD = 89\text{mm.s}^{-1}$ ),  $F(1,18) = 241.355$ ,  $p < .001$ ,  $\eta^2_p = .93$ , and  $MT_2$  was shorter at Fast speed ( $M = 438$ ,  $SD = 61\text{ms}$ ) than at Preferred speed ( $M = 742$ ,  $SD = 101\text{ms}$ ),  $F(1,18) = 280.882$ ,  $p < .001$ ,  $\eta^2_p = .94$ . A global effect of Condition was also observed on  $APV_2$ ,  $F(1,18) = 5.831$ ,  $p = .026$ ,  $\eta^2_p = .24$ , and  $MT_2$ ,  $F(1,18) = 15.251$ ,  $p = .001$ ,  $\eta^2_p = .46$ , indicating that when placing under noPause condition, the peak velocity was higher and the movement duration was shorter. More importantly, no interaction effects were found on either  $APV_2$ ,  $F(1,18) = 4.072$ ,  $p = .059$  or  $MT_2$ ,  $F(1,18) = 1.536$ ,  $p = .231$ . Figure 6 presents the group results for both the pick and place kinematic parameters of the motor sequence.





**Figure 6.** APV<sub>1</sub>, APV<sub>2</sub> and End Position Errors - Group results (*Experiment 3*).

Means and standard error for APV<sub>1</sub> and APV<sub>2</sub> during noPause (A) or Pause (B) conditions. Dark-grey : Preferred speed White-grey : Fast Speed. Note that for the placing element, no effect of Pause is found, however for the reaching element, at fast speed peak velocity is higher under noPause ( $M = 1420$ ,  $SD = 223\text{mm.s}^{-1}$ ) than under Pause condition ( $M = 1168$ ,  $SD = 170\text{mm.s}^{-1}$ ). Moreover, under Pause, peak velocity of reaching element at fast speed is not different than at preferred speed ( $p = .40$ ).

## Correlation analysis between elements 1 and 2.

**MT<sub>1</sub>-MT<sub>2</sub>.** Results showed a significant global effect of Condition on the coefficient of determination ( $R^2$ ) of the linear regression between MT<sub>1</sub> and MT<sub>2</sub>. The  $R^2$  was higher under noPause condition ( $M = 0.253$ ,  $SD = 0.200$ ) than under Pause condition ( $M = 0.127$ ,  $SD = 0.120$ ),  $F(1, 18) = 5.637$ ,  $p = .029$ ,  $\eta_p^2 = .24$ . The effect of Speed on the correlation strength was none significant,  $F(1, 18) = 0.865$ ,  $p = .365$ . The interaction between Condition and Speed was found to be close to significance,  $F(1, 18) = 3.969$ ,  $p = .062$ . Post-hoc Scheffé tests confirmed that the Condition effect between noPause ( $M = 0.283$ ,  $SD = 0.274$ ) and Pause ( $M = 0.063$ ,  $SD = 0.072$ ) was only significant at Fast Speed,  $p = .032$ . No

such effects were observed at Preferred Speed between noPause ( $M = 0.223$ ,  $SD = 0.220$ ) and With Pause ( $M = 0.190$ ,  $SD = 0.219$ ),  $p = .969$ . Figure 7 illustrates these results in a typical individual (subject LO) performing under Fast Speed condition.

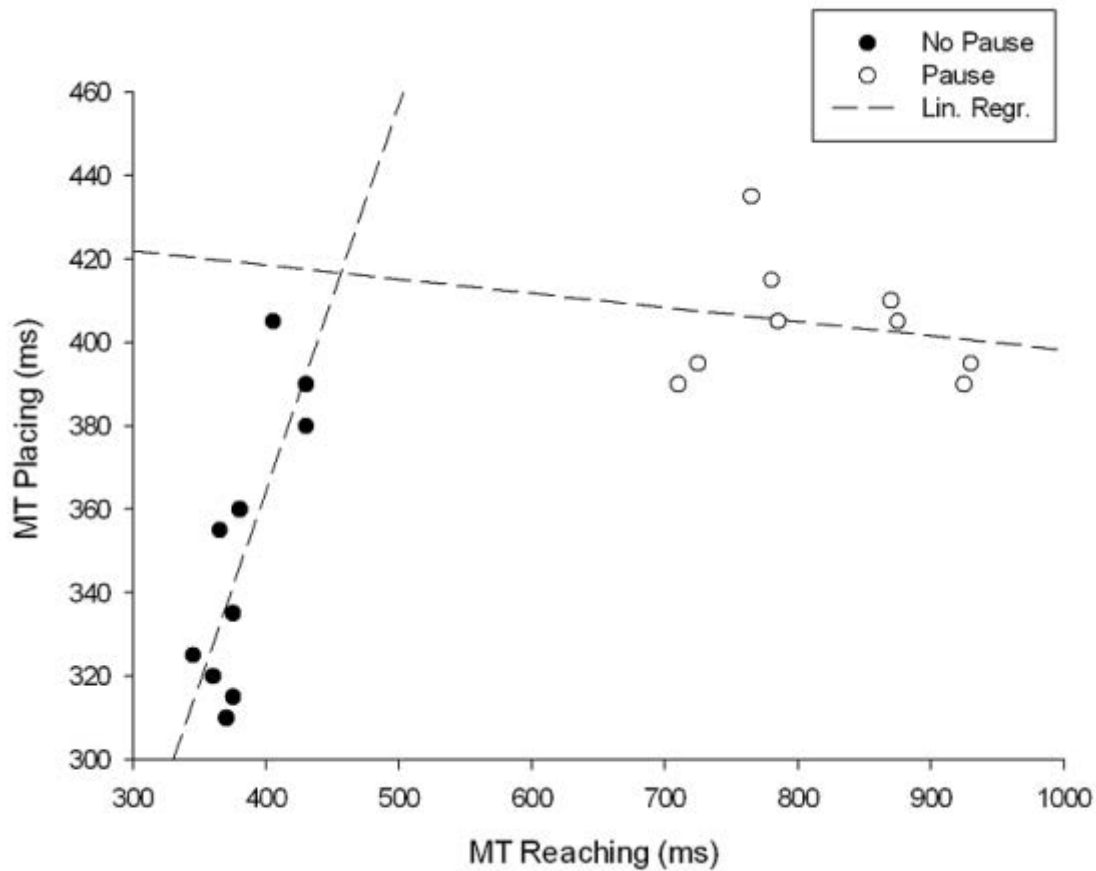


Figure 7. Scatterplots of  $MT_1$  and  $MT_2$  under noPause and Pause conditions (*Experiment 3*).

Linear regression of movement times (MT) for placing movements performed at fast speed. White and black filled circles are individual trials for No pause and Pause conditions, respectively. Dashed lines represent simple linear regressions lines (least squares fit). Group results showed significant differences between coefficients of determination under noPause and Pause conditions (noPause:  $r^2 = 0.24 \pm 0.05$  vs. Pause :  $r^2 = 0.07 \pm 0.02$ ,  $p = .011$ )

**APV<sub>1</sub>-APV<sub>2</sub>.** There were no significant effects of Condition on the coefficient of determination ( $R^2$ ) of the linear regression between  $MT_1$  and  $MT_2$ ,  $F(1, 18) = 3.072$ ,  $p = .097$ ,  $\eta_p^2 = .15$ . However, a global effect of Speed was found, indicating that the peaks of velocity within the same sequence were better correlated when performed at Fast Speed ( $M = 0.209$ ,

SD = 0.133) than at Preferred Speed ( $M = 0.137$ ,  $SD = 0.081$ ), regardless of the Condition,  $F(1, 18) = 0.865$ ,  $p = .365$ . More importantly, the interaction between Condition and Speed was significant,  $F(1, 18) = 4.316$ ,  $p = .052$ . The post-hoc Scheffé tests confirmed that the Condition effect between noPause ( $M = 0.293$ ,  $SD = 0.220$ ) and Pause ( $M = 0.126$ ,  $SD = 0.037$ ) was only significant at Fast Speed,  $p = .042$ . No such effects were revealed between noPause ( $M = 0.143$ ,  $SD = 0.164$ ) and Pause ( $M = 0.130$ ,  $SD = 0.151$ ) at Preferred Speed,  $p = .996$ .

**Conclusions.** The results presented in this third experiment confirmed the results obtained in the second one: when the sequence is continuous (noPause), the constraints of the second element are back propagated to affect the execution properties of the first element of the sequence. This is yet again the demonstration of coupled planning in an ecological pick to place object manipulation task. When the sequence was interrupted, coupled planning was prevented: the two elements were prepared and executed as independent movements, with correlations that were no longer significant between movement times and peak velocities. Interrupting the sequence at preferred speed had no impact, confirming the fact that the absence of time pressure (low task difficulty), the planning strategies are different in nature. This reinforces the idea that at preferred speed, the two movements can be planned separately in a step-by-step strategy. Conversely, at fast speeds, participants relied on a global anticipatory strategy (coupled planning) maybe to insure correct performance outcome (maintain the object stable within grip); this global planning is nevertheless possible only when all parameters and constraints – both in space and in time – can be predicted allowing for early encapsulation of the motor elements that constitute the object directed sequence.

## GENERAL DISCUSSION

During face to face interaction, one can usually read a person's intentions are, especially if that person is a close friend. A quick glimpse at that person's actions can usually give a good insight on what the next move will be, e.g., *he is reaching to give me his keys of the car*. This quick reading of intention is common in human behaviour but hard to implement in an artificial system. The idea developed in the present manuscript is that intention reading is possible because we are able to implicitly extract those slight and relevant kinematic variants that occur early in their motor sequence, in order to anticipate the intended motor goal.

Scientific evidence is showing that adults can read intentions through simple observation even when little or no contextual information is available (Becchio et al., 2012; Manera et al., 2011; Sartori et al., 2011), suggesting that early variants in movement kinematics must be present both in function of the nature of motor goal but also of the context in which the sequential action is performed. In the present study, we demonstrated that when anticipatory strategies are possible, accuracy constraints back propagate and influence motor kinematics very early on during the execution of the very first element of the motor sequence. These results were obtained without modifying the nature of the task. It replicates previous data obtained in sequential pointing, and importantly confirms the presence of back propagation laws in object pick and place tasks. We furthermore revealed the importance of task difficulty and predictability for the use of coupled planning strategies: with unpredictable time breaks, global planning could not be implemented; with low task difficulty, global planning was not implemented. In the following sections, these findings are further described and discussed relative to their implication for the development of humanoid robots that may offer intuitive interaction if correctly implemented with coupled planning laws of movement.

**Coupled Planning in Sequential Actions.** In this study, we addressed the question whether object manipulation sequences can be planned globally. Coupled planning is basic and essential for the performance of fast complex sequential motor actions (Adam et al., 1995; Fischman & Reeve, 1992; Rand et al., 1997; Rand & Stelmach, 2000; Short, Fischman, & Wang, 1996; Sidaway, Sekiya, & Fairweather, 1995) and especially when participants are required to also obey a high accuracy constraint (Ketcham, Seidler, Van Gemmert, & Stelmach, 2002; MacKenzie, Marteniuk, Dugas, Liske, & Eickmeier, 1987; Rand & Stelmach, 2000; P. Weiss et al., 1997). In object manipulation experiments, a large number of studies have reported evidence of coupled planning depending on the final end-state comfort position for the arm (David A. Rosenbaum & Jorgensen, 1992; Short & Cauraugh, 1997), on the grasp orientation and grasp height (Cohen & Rosenbaum, 2004; for a review see: Rosenbaum, Chapman, Weigelt, Weiss, & van der Wel, 2012) as well as the second action type (Armbrüster & Spijkers, 2006; Johnson-Frey et al., 2004; Marteniuk et al., 1987; Mason, 2007). Despite the large number of empirical results suggesting early modulations of the reach-to-grasp (first) element due to imposed constraints on the second element of a motor sequence in the pointing literature, striking evidences have not yet been presented for object grasping. Ansuini and al. in 2006 reported an elegant study showing that the very same object placed in either a low accuracy or a high accuracy niche is grasped in contrasting ways. However, in this study, because the distance for the second element (6cm) was not similar to the first element (30cm) the authors could not compare directly the cross-correlation proprieties of the two motor elements. Moreover, because the niche was a 3D object for which the shape was manipulated along with its size, the effect found in the first motor element could have been due to a difference in second action type and not strictly because of its size. The authors hypothesized that subjects may have adopted the strategy of an early shaping of the hand to bypass the incongruent shape information provided by the nearby low accuracy niche. In contrast, when the niche had the same shape as the object (high accuracy niche), the lack of potential conflict between the shape of the niche and the shape of the target object may

have allowed for a gradual on-line hand shaping similar to that found for the no-niche condition (Ansuini et al., 2006). Nonetheless, some of the most recent studies (Johnson-Frey et al., 2004; Mason, 2007) have reported contradicting results for the accuracy hypothesis (Short & Cauraugh, 1999), thus leading to confused conclusions about the use of coupled planning during object pick and place tasks.

In the present study, we reported data demonstrating that - without any modification of the second movement type or target shape, the final accuracy constraints are fully taken into account prior to the initiation of the first segment of the sequence. This strongly argues in favor of the use of the global planning strategy in object pick and place actions. In the following, we discuss our results in detail and will conclude that global planning is indeed present for sequential manipulation object but only if both time and space constraints of the task are fully predictable by the agent as early as the planning phase of the motor sequence.

**Effects of end-point constraints.** In the first experiment, subjects were asked to reach and grasp a wooden cylinder and place it on a place pad that could be of 2 different sizes. Two groups of subjects participated in the experiment: one group was given a verbal “go” to initiate their movement; the other group reacted to light-emitting diodes. This latter group also benefited from an automatized reward system based on real-time analysis of final object position that told them whether the second place movement had been correctly performed (speed and accuracy). The results were contrasted. Overall, in the verbal instruction group, we observed lower speed, smaller end position errors and longer movement times than for the real-time group. No effects of final accuracy constraints were found on the first pick movement of the sequence, even in those trials performed as fast as possible; some participants remarked on the length of the session duration with beginning of upper arm muscle fatigue. On the contrary, those participants benefiting from the real-time controlled system expressed the feeling that time went by fast; with a pleasurable experience even through the task was indeed repetitive. As in the other group, no effects of terminal accuracy were observed on the first pick movement of the sequence but for those trials performed at

preferred speed only. Indeed, when performing the second place movement as fast as possible, results revealed a strong speed-accuracy trade off, coupling those constraints set upon the second place element with the speed of the first pick element of the sequence. More specifically, under fast speed and high accuracy constraints, participants showed longer movement times and lower peaks of velocity than under low accuracy constraints trials. These results suggest coupled planning of the two motor elements within a unique motor plan. However, because the results were leading us to a different conclusion than previously reported (Johnson-Frey 2004, Mason, 2007), we conducted a second experiment to confirm the findings.

To do so, we replicated the experimental setup using the real-time protocol only with an additional third intermediate accuracy constraint. We also conducted a more in-depth statistical analysis of the results. The data from this second experiment confirmed those obtained in the first: terminal accuracy constraints are taken into account during the initial planning phase of the sequence and thus, the terminal constraints back propagate to modify the kinematic patterns of the very first motor element. More specifically, for the reaching movement we found under fast speed and high accuracy constraints longer movement times, lower peaks of velocity and lower peaks of acceleration than under fast speed and mid or low accuracy pads. For the grasping component, we observed higher maximum grip apertures, shorter times to maximum grip aperture and higher velocity peaks of grip aperture under fast speed as compared to preferred speed. Taken together, the results presented in experiments 1 and 2 suggest that planning a sequence of two motor elements back propagates terminal constraints. These results confirm those reported in the pointing literature. In the two studies previously mentioned for which no effects of final accuracy constraints were reported (Johnson-Frey et al., 2004; Mason, 2007), the first motor element of the sequence was in fact sensitive to the nature of the second element. The authors were hence in favour of a global planning hypothesis and suggested that actions must be determined both by perceived spatial demands of immediate movement and the intended goal of action sequence. Our results are in

line with this hypothesis but further demonstrate that the spatial demands of the second target can quickly be taken into account if relevant enough to insure final goal performance. Our statistical analysis further suggests that not only global planning but *coupled planning* occurs when both motor elements are specified in advance both through space and time.

The results of the second experiment are the first to confirm that the *accuracy* constraints of the end position are back propagated onto the early reach segment of the sequential movement. Without modifying any final goal position, orientation, initial instructions, target shape, context or second action type, we demonstrate early kinematic modulations relative only to the spatial and the temporal demands of the task. In addition, correlation analyses also confirmed that the parameters of the second and the first elements showed a very tight relationship under fast speed and accuracy constraints. These results confirm the *coupled planning* of both first and second element together, and show that the constraints that are imposed either at a global (speed) or a more local level (accuracy) shape the kinematic patterns of the entire sequence and thus, may be used very early by an external observer to read motor intention. In contrast to speeded constraints, when moving at preferred speed, the kinematics of the two elements showed low levels of correlations suggesting that participants were using different planning strategies when moving slowly than when moving fast, i.e., using step-by-step planning and coupled planning, respectively.

**Unpredictability to un-couple sequential actions.** In the third experiment, we addressed the question of the timing predictability required to induce the coupled planning strategy. Here, we asked participants to reach and grasp the object but either to wait for a second tone before lifting and placing the object or to perform the sequence without any interruption. By adding a random time interval between the first and the second motor element, we were able to prevent the back propagation effect of the speed accuracy trade-off induced in the second element. More specifically, our data showed that peak velocity was higher and reaching duration was shorter during noPause conditions than during Pause conditions. When the sequence was interrupted, there was also an absence of effects of the speed instruction of the second place



element on the first pick element of the sequence. Moreover, the correlations between the two motor elements were strongly weakened by the presence of the unpredictable time break. As expected, *coupled planning* only occurred under fast speed and in absence of time break; and in such case, results demonstrated high levels of correlation between the kinematic parameters of the first and the second motor elements, replicating those results from study 1 and 2. Hence, we concluded that the effect of the unpredictable time break strongly impaired the global cognitive planning strategy and lead to a decoupling of the elements that constituted the motor sequence.

These results suggest that when the timing of the entire sequence is not predictable, participants are more likely to use a step-by-step strategy. They first plan and execute the first motor element, then plan and prepare the second motor element but inhibit its initiation until the second tone is heard. Thus, the high level of linkage between the two movements disappears. However, when the agent can fully anticipate the temporal constraints, the more efficient *coupled planning* strategy is implemented and both motor elements are affected by internal (intentions) and external (contexts) parameters. It is the case that the step-by-step strategy would more often lead to low fluidity between motor elements with more time spent between the end of one element and the start of the next, as what is typically seen in patients suffering from schizophrenia (Delevoye-Turrell, Giersch, Wing, & Danion, 2007). Such actions may be less costly in cognitive efforts (Delevoye-Turrell & Wing, 2004), but do not afford the fluency that enable efficient social interaction (Hayes, Paul, Beuger, & Tipper, 2008). The following paragraphs consider the problem of anticipatory planning for optimal social interactions and especially the functional relevancy of back propagation strategies for reading motor intentions when interacting with humanoid robots.

**Perspectives for intuitive interactions with social robots.** The current computational models of motor control are based on optimized function costs that are often named minimum-X, e.g., jerk, torque change, energy, time, variance (Flash & Hogan, 1985; Shadmehr & Wise, 2005; Wolpert et al., 2003; for a review see: Todorov, 2004). The findings

that the movement kinematics are affected by both the timing and the spatial accuracy constraints set upon the terminal goal is in line with such optimization functions. This would have been the result of a process of selection of relevant action features, which would ensure an executive plant with minimum end-point variances even at the fastest execution speed. However, even the most advanced optimization model does not explain why, in a 2-element sequence, the kinematics of the first pick element would be influenced by the final target constraints affecting the second place element, even though all other parameters are maintained strictly identical (object distance, object size, etc...). A simple criterion based on initial position, object size and intermediate object position would have predicted similar movement patterns in the first pick element under all experimental conditions. Our results demonstrated that this is not the case for human arm movements and thus, our data reinforces the need in computational neurosciences to take into account contextual and distal constraints in addition to the other classic single-step more localized constraints (e.g. minimum-X) during the planning and execution of complex motor sequences.

The results we found in experiment 1 also emphasize the importance of the constraints to be set by the environment in contrast to that imposed through verbal instructions. Indeed, our participants were able to learn implicitly the speed-accuracy trade-off problem, and to adapt behaviour by repeatedly correcting their motor output on the basis of both online sensory feedback (visual and proprioceptive) but especially the reliable and accurate feedback provided by the interactive setup. The solution adopted by the brain after years of motor learning is a cognitive solution for which coupled planning is applied to optimize motor fluency and accuracy through the back-propagation of terminal constraints on early phase of action preparation. Very few computational models are able to conceptualize this ability of the motor system, especially for complex sequences requiring e.g. reach, grip, lift, transport and safe placing (Oliver Herbort & Butz, 2011). In fact, in most computational models, sequencing motor actions is often attributed to basal ganglia. However, the classical basal ganglia models do not specify the “content” of action (Doya, 2000) and only deal with the

timing problem of motor sequencing. In this matter, cognitive models are now well able to explain how the brain may be able to implement and integrate contextual constraining features for fluent and efficient motor interactions.

**Encapsulating motor elements together for motor fluency.** During motor planning, the process of action-features selection involves at least two cognitive stages (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz & Hommel, 2002; Stoet & Hommel, 1999). First, the relevant body features for a given task are activated. Second, these features are presumably bound together into a composite representation of motor and perceptual specifications (Stoet & Hommel, 1999). The binding procedure is a shielding process that preserves action representations from interferences with other cognitive process (thought distractors) and elements present in the workspace (physical distractors). Through this encapsulating process, the goal representation remains intact throughout the course of action preparation and execution. This binding of features has been referred to in the literature as “occupation” (Schubö, Prinz, & Aschersleben, 2004), “encapsulation” (Müsseler, 1999) or “a common event file” (Hommel, 2004). The data presented here suggests a encapsulating procedure of those kinematic features coding for both motor elements in ecological object pick and place tasks. This is in favour of a common code of perception and action, which would link perceptual and motor representations into a single cognitive structure, i.e., in an integrated, task-tuned network of feature codes of perceived and to-be-produced events (Hommel et al., 2001).

The correlation data that is presented in study 2 further argues in favour of a unique representation in object manipulation tasks. Indeed, when the constraints are applied locally (terminal accuracy) and globally (speed), both motor elements are affected in a similar fashion, and at the level of the same motor parameter. Thus, the coupled planning would be a selected mechanism that would modulate a given parameter of a global plan. This strategy has definite advantages as compared to step-by-step strategies. First, coupled planning would provide the means to ensure the fluency of execution of a motor sequence. If both motor

elements are planned before the execution and are characterized by similar velocity and acceleration patterns, the overall action will possess smoother transitions and lower biomechanical pressure on muscles and joints at the critical moment of passage from one element to the next. When planning to execute a sequence of two motor elements, using predictive models of both motor elements would also have a learning advantage. For example, the comparison of the predictions based on motor commands and the current state of the body would lead to fast and rapid adaptations at any point of the sequence when an error is detected. The predictive models could also be used to select relevant features for both motor elements and check the compatibility between sub-elements, to ensure successful outcome. Most importantly maybe, coupled planning would allow for an external observer to read the initial intention by detecting the slight variants in kinematics of the early reach-to-grasp movement – compared to the kinematics that would be expected from the optimal strategy (Lewkowicz et al. 2013). It has been shown for example that for the observers, the use of internal models based on simulated motor commands of the first motor element would help to understand and predict the final target of the sequence. Such predictions could serve to stabilize the ongoing percept by assisting perceptual disambiguation (Wilson & Knoblich, 2005). Thus, coupled planning would have an important role to play in successful social interaction processes, and would be able to give the observer access to intention reading (Becchio, Sartori, Bulgheroni, & Castiello, 2008a; Manera et al., 2010; Sartori et al., 2011). In the context of a national French grant, we are now in the process of implementing artificial neural networks with kinematic decoding capacities in order to test the hypothesis that (1) coupled planning can be implemented for different types of sequential actions (whole body movement and two handed object manipulation) and that (2) the detection of early deviants in kinematic patterns can help trigger intuitive interactions in both human-human and human-robot dyads.

## CONCLUSIONS

In the present contribution, we report data that demonstrate a couple planning of kinematic parameters between multiple elements of a pick and place motor sequence. When all constraints were fully predictable, i.e., timing, space and accuracy constraints, we showed that the constraints set on the final element back propagated to influence the very early phase of the first element of the sequence. This coupled planning strategy is in line with the cognitive encapsulating hypothesis of motor and perceptual features that can provide action fluency, efficiency and adaptability compared to those behaviors observed when a more basic step-by-step planning strategy is used. Coupled planning would also be important because it modifies kinematic patterns in a systematic way that become meaningful for external observers by providing valuable non-verbal social cues of motor intention. If reading an agent's intention is possible with early kinematic variants, then coupled planning would be a way to allow for fast understanding and adaptations for intuitive social interactions. As such, coupled planning laws need to be integrated in those computational and pathological models of motor control that are being used to create interactive humanoid robotic systems.

## **3.2. Effet de la variabilité terminale sur la vitesse d'exécution et le couplage des mouvements inclus dans une même séquence motrice.**

### **Résumés**

Dans la littérature du contrôle moteur, plusieurs théories récentes proposent que le cerveau agit comme un contrôleur optimal, en planifiant les actions motrices de sorte que la variabilité d'exécution soit minimale (Harris & Wolpert, 1998; Todorov, 2004; van Beers, Haggard, & Wolpert, 2004). Dans la présente contribution, nous avons étudié l'effet d'une instabilité terminale sur les caractéristiques de planification précoce dans une tâche séquentielle de manipulation d'objet. La tâche du sujet était de saisir un objet pour le placer sur un réceptacle dont la stabilité était plus ou moins prévisible. Les résultats ont révélé une modification précoce de l'exécution motrice avec notamment une réduction significative de la vitesse d'exécution dès le premier élément de la séquence, en réponse à une instabilité de la plateforme terminale. De plus, un meilleur couplage a été observé entre les paramètres cinématiques des sous-éléments de la séquence d'autant plus que l'instabilité induite expérimentalement était prévisible. Ces résultats sont en accord avec le modèle de planification par anticipation (Fischer, Rosenbaum, & Vaughan, 1997; O. Herbort & Butz, 2007) et apportent un éclairage nouveau sur la notion de rétro-propagation des contraintes environnementales et également, sur le phénomène de planification couplée.

**Mots-clés :** action, planification, couplage, variabilité, séquence, rétro-propagation.

In the field of motor control, recent theories have described the brain as an optimal controller, which would plan motor actions in such a way to minimise execution variability (Harris & Wolpert, 1998; van Beers, Haggard, & Wolpert, 2004 for a review see : Todorov, 2004). In the present contribution, we investigated the effects of terminal instability on the early components of action planning in the sequential task of pick and place. The subjects' task was to reach for an object and to place it on a moving platform. The instability of this final pad was more or less predictable. Results showed that terminal variability modulated action execution with significant kinematic changes as early as the first element of the sequence. In addition, greater coupling between the different elements of the sequence was revealed especially for those conditions in which the platform's instability could be predicted. These results are in line with the model of predictive global planning (Fischer, Rosenbaum, & Vaughan, 1997; Herbolt & Butz, 2007), and bring new insights on the phenomena of coupled planning in relation to context instability.

**Keywords:** action, planning, coupling, kinematics, reaching, grasping, placing, sequential, variability, back-propagation.

## **Introduction**

Une des caractéristiques prédominantes du contrôle moteur chez l'humain est la variabilité de nos mouvements qui limite la précision avec laquelle nous pouvons réaliser une action. Par exemple, si on demande à quelqu'un de répéter plusieurs fois le même geste (même effecteur, même point d'origine, même point d'arrivée), on observe rapidement une certaine variabilité dans les caractéristiques cinématiques de ces mouvements (vitesse, accélération, durée, précision terminale). Prenons le cas d'un mouvement réalisé de la

manière la plus rapide possible. Dans ce cas, les forces de contraction musculaires sont plus importantes que celles nécessaires pour l'exécution d'un mouvement lent. Il a aussi été montré que la force du signal neuronal qui est envoyé à ces muscles est nécessairement plus importante pour obtenir des contractions rapides et puissantes. Cependant, en augmentant la force du signal, le bruit est aussi augmenté. Harris & Wolpert (1998) suggèrent alors que ce bruit « neuronal » serait à l'origine de la faible précision terminale observée lors de mouvements très rapides. A l'inverse, si pour une tâche donnée, les contraintes de précision sont plus fortes, alors la durée des mouvements pour la réaliser sera plus grande (Loi de Fitts). Par exemple, passer un fil dans le chat d'une aiguille demande beaucoup plus de temps que de passer un lacet dans le trou d'une chaussure. C'est ainsi qu'il a été suggéré que le paramètre de contrôle de nos mouvements serait leur *variabilité*. Afin de minimiser les conséquences délétères du bruit, le cerveau chercherait à réduire l'imprécision de nos gestes en adoptant des stratégies permettant de contrôler la variabilité des mouvements (Harris & Wolpert, 1998). Néanmoins, l'origine de cette variabilité est multiple et de nombreuses études expérimentales ont ainsi été menées pour identifier quelles sont les sources de variabilité dans les différentes étapes du contrôle moteur.

Lorsque nous souhaitons atteindre une cible avec notre main, au moins trois étapes sont nécessaires pour nous permettre de produire le mouvement : la localisation, la planification et l'exécution. Etant donné qu'elles s'effectuent à travers le système nerveux, chacune de ces étapes est sujette au bruit inhérent à la transmission neuronale. Dès lors, nous pouvons légitimement nous demander quelles sont les conséquences de ce bruit dans leur réalisation.

## **Localisation**

Lors de l'étape de localisation, les positions de la cible et de la main sont obtenues à l'aide d'informations sensorielles à la fois visuelles et proprioceptives. La question de la variabilité pour l'étape de localisation a fait l'objet de nombreux travaux expérimentaux. Par



exemple, Rossetti et collaborateurs (1994) ont demandé à des sujets de pointer en direction d'une cible selon différentes conditions de visibilité. Les sujets pouvaient voir ou non leur main préalablement à l'initiation du mouvement mais jamais pendant son exécution. Leurs résultats montrent que la variabilité terminale des actions était significativement réduite lorsqu'à la fois la main et la cible étaient visibles, par rapport à une condition où uniquement la cible est visible. Les auteurs en ont conclu que les informations visuelles de position *initiale* de la main et de la cible sont importantes afin de compenser la variabilité *initiale* associée à la programmation du geste, et que ce phénomène ne dépendait donc pas du retour visuel pendant l'exécution de la tâche (Rossetti, Stelmach, Desmurget, Prablanc, & Jeannerod, 1994). D'autres auteurs ont trouvé des résultats similaires en manipulant la position de la cible en 2D (Van Beers, Sittig, & Denier van der Gon, 1998) ou en 3D (McIntyre, Stratta, & Lacquaniti, 1997). Dans cette dernière étude, un délai temporel était manipulé entre la présentation de la cible et l'initiation du mouvement. Les résultats ont montré qu'en augmentant ce délai, l'erreur terminale était très largement accrue, et ce, dans les trois dimensions de l'espace. Ces auteurs ont alors suggéré que l'information de position de la cible était maintenue en mémoire dans une forme dont la stabilité est indépendante de la direction de l'espace (McIntyre et al., 1997).

Ces études ont permis de mettre en évidence des effets de la variabilité liée à l'étape de *localisation*. Afin de réduire cette variabilité, le cerveau doit avoir accès à des informations visuelles précises pour reconstruire l'espace en trois dimensions dans lequel le mouvement doit être exécuté. Une fois déterminées, les informations de localisation vont ensuite être utilisées lors de l'étape de planification du mouvement. La façon le cerveau va alors contrôler la variabilité inhérente à l'étape de planification est aussi une question qui a suscité de nombreux travaux expérimentaux.

## Planification

La planification du mouvement est l'étape qui consiste en la sélection des commandes motrices qui vont permettre de produire un mouvement désiré, de la position initiale jusqu'à la position finale. Cette étape utilise les sorties de l'étape de localisation comme entrée et génère des commandes motrices en sortie. Par exemple, pour atteindre et saisir un objet, nous devons à la fois déplacer notre main en direction de cet objet (composante de transport) et ouvrir notre main de façon à l'attraper correctement (composante de manipulation). En demandant à des sujets de saisir des objets placés à différentes distances, Jeannerod (1984) a analysé les caractéristiques de la coordination spatio-temporelle entre les composantes de transport et de manipulation. Les résultats ont suggéré que le pic d'ouverture maximal de la pince digitale est hautement corrélé avec le pic de décélération du transport de la main (à environ 75% de la trajectoire totale). De plus, ce lien existe même dans les conditions pour lesquelles les sujets n'ont aucun retour visuel de leur membre en mouvement (M. Jeannerod, 1984). Les théories motrices actuelles concèdent donc la possibilité qu'une programmation centrale, commune et globale, soit mise en place par le cerveau pour coordonner les deux composantes d'une même séquence de mouvement. Au cours de cette étape de planification, les commandes motrices pour le bras et la main seraient déterminées préalablement à l'initiation du mouvement de préhension afin de faciliter la probabilité d'atteinte de l'objectif moteur (prendre l'objet), de part notamment la maîtrise de la variabilité inhérente aux différentes boucles de contrôle impliquées pour l'exécution de séries d'actions motrices (Kawato, 1999).

Pour considérer le rôle des boucles de retour sensoriel dans l'optimisation des mouvements de saisie, la question de la compensation en absence de vision a été étudiée par Wing et collaborateurs (1986). Leurs études ont montré que la taille de l'ouverture de la pince digitale pouvait compenser les facteurs qui influencent la variabilité liée à l'exécution de la composante de transport. En effet, pour un mouvement réalisé le plus rapidement possible, l'ouverture de la pince est plus grande comparativement à ce qui est observé dans le cas d'un

mouvement réalisé à vitesse naturelle. De plus, l'ouverture maximale de la main est encore plus grande lorsque ce mouvement est réalisé les yeux fermés ou sans retour visuel. Dans les deux cas, le geste présente des erreurs spatiales significatives. Wing et collaborateurs ont ainsi suggéré qu'en anticipation des erreurs de planification, l'ouverture de la pince digitale est programmée de façon à compenser l'imprécision de la composante de transport (Wing, Turton, & Fraser, 1986).

Enfin, des sources de variabilité additionnelles ont pu être identifiées, lors de l'étape de planification, en prenant en compte par exemple les propriétés fonctionnelles de l'espace d'action (direction et amplitude). Plus spécifiquement, Gordon et Ghez (1994) ont demandé à des sujets de pointer par des mouvements rapides et répétés des cibles dans plusieurs directions de l'espace. Par l'analyse de la distribution spatiale des points terminaux, ils ont alors montré que les distributions étaient elliptiques et orientées selon la direction des mouvements exécutés. Selon les auteurs, ces résultats seraient liés au fait que les mouvements sont planifiés dans un système de coordonnées dont l'origine se situe au niveau de la position initiale de la main, selon un codage vectoriel du mouvement entre les points d'origine et d'arrivée. Par ailleurs, leurs résultats montrent que les erreurs de pointage dans l'axe du mouvement augmentent avec la distance entre le point d'origine et la cible alors que l'erreur de direction (axe perpendiculaire à l'axe du mouvement) est globalement constante, d'où la forme elliptique de la distribution des points terminaux. Pour ces auteurs, il semblerait alors que la variabilité directionnelle soit indépendante de la variabilité liée à l'amplitude du mouvement ; ces deux éléments seraient alors planifiés séparément (Gordon, Ghilardi, & Ghez, 1994). En résumé, lors de l'étape de planification, les sources de variabilité vont émerger notamment lors des calculs de transformation de coordonnées entre l'espace visuel (œil-centré), et l'espace articulaire (corps-centré) avec des erreurs liées au codage vectoriel de l'espace (Vindras & Viviani, 1998). Toutefois, une fois que la sélection des commandes a été effectuée, le « programme moteur » est utilisé pour exécuter l'action. De nouveau, au cours de cette dernière étape d'exécution, d'autres sources de variabilité peuvent aussi intervenir.

## Execution

Lors de l'étape d'exécution, les commandes motrices sont envoyées aux muscles de façon à ce que le mouvement soit effectué. La variabilité observée dans le mouvement pourrait ainsi refléter le bruit du signal neuronal reçu par les organes terminaux. Pour tester cette hypothèse, des sujets ont eu pour tâche de répéter un grand nombre de fois les mêmes mouvements de pointage pour qu'après apprentissage, la variabilité liée aux étapes de localisation ou de planification soit minimisée. En augmentant la quantité de répétitions dans un design d'expérience en bloc, Van Beers, Haggard et Wolpert (2004) ont ainsi montré que la variabilité liée à l'exécution n'était pas uniforme dans toutes les directions de l'espace. En effet, la distribution des points terminaux ne se répartissait pas forcément selon l'axe principal entre la position initiale et la position finale contrairement au modèle de codage vectoriel de Gordon et Ghez (1994). Les résultats montrent que la distribution des points terminaux se répartissait selon la *courbure* de la trajectoire suivie par la main (van Beers et al., 2004). Cette courbure dépendrait à la fois de la vitesse du mouvement et de la distance à parcourir, et son origine pourrait être liée aux propriétés biomécaniques de l'effecteur (muscles et ligaments du bras).

En résumé, les études rapidement décrites ci-dessus ont permis de mettre en évidence les différentes sources de variabilités qui interviennent à chaque étape du contrôle moteur, de la localisation de la cible jusqu'à l'exécution du mouvement. Pour chacune d'entre elles, le système de contrôle moteur humain s'organise de façon à minimiser les effets délétères de cette variabilité. Cependant, la plupart de ces études ont été faites dans un environnement stable et prédictible, avec des contraintes qui sont connues très tôt pendant la phase de planification. Tout un autre groupe de chercheurs se sont posé alors la question de mieux déterminer la part de variabilité motrice qui serait liée à une instabilité de l'environnement et notamment, suite à une modification des propriétés de l'espace d'action dans lequel nous effectuons les mouvements.

## **Perturbations mécaniques de l'environnement**

Pour manipuler la prédictibilité des interactions entre un acteur et son environnement proche, différents protocoles expérimentaux ont été développés. Par exemple, certains auteurs se sont intéressés à la manière dont les sujets humains étaient capables d'adapter leurs réponses suite à une perturbation mécanique. Dans ces études, les effets d'une perturbation au cours de l'exécution ont été étudiées en mettant les sujets dans une chambre rotative (Lackner & Dizio, 1994) ou encore, en utilisant des champs de forces simulés via des manipulandum robotiques (Shadmehr & Mussa-Ivaldi, 1994; Won & Hogan, 1995). Ces études montrent que les sujets vont compenser les perturbations en modifiant l'impédance musculaire, c'est-à-dire les relations entre la force du muscle et son étirement induit. L'équipe de Shadmehr et collaborateurs ont ainsi montré que l'adaptation comportementale se réalise via la formation d'un modèle interne approprié pour compenser la variabilité qui émerge des interactions successives avec l'environnement. Cette adaptation est particulièrement visible lorsque le champ de force est supprimé soudainement et que les sujets continuent à exécuter les mouvements sans retour visuel de leur effecteur. Les résultats révèlent que les déviations par rapport aux trajectoires sont alors toujours présentes même après 250 essais sans champ de force. De plus, cette adaptation se transfère d'une direction de l'espace à une autre (Shadmehr & Mussa-Ivaldi, 1994). Ces résultats sont des arguments solides en faveur d'un modèle interne de planification du mouvement, établi préalablement à l'exécution, qui peut être modifié par les interactions successives du sujet avec son environnement.

Une autre approche, beaucoup moins développée, a été de manipuler la stabilité *terminale* de l'environnement, i.e., des propriétés de l'espace d'action dans lequel se place l'objectif moteur. Dans le cas d'une manipulation d'objet, la question de la stabilité du réceptacle sur lequel un objet est posé a été développée dans une étude menée par Klatzky et collaborateurs (1995). Les sujets avaient pour tâche de prendre des objets qui étaient sur une plateforme glissante ou non-glissante. Les résultats ont montré que la durée des mouvements pour atteindre l'objet était significativement plus grande lorsque ce dernier était placé sur un

réceptacle glissant. Les auteurs suggèrent alors que la stabilité de la main au contact de l'objet est prise en compte lors de la planification en tant que contrainte supplémentaire (Klatzky, Fikes, & Pellegrino, 1995).

Toutefois, aucune étude à ce jour n'a posé la question de l'instabilité de l'environnement, non pas au contact de l'objet mais lors du contact de cet objet sur la surface sur lequel on va le poser, c'est-à-dire à la fin d'une séquence de deux sous-éléments moteurs. Dans une tâche de saisie, quelle serait alors l'influence de l'instabilité du réceptacle sur la préparation et la production de la séquence d'action ? Dans la présente étude, nous proposons de tester ces effets dans une situation de capture de mouvement interactive au cours de laquelle les sujets avaient pour instruction de saisir puis de déplacer un objet d'une plateforme stable à une plateforme instable. De plus, nous avons réalisé une situation interactive au cours de laquelle les sujets recevaient des renforcements en temps réel sur leurs propres performances. Ainsi cela nous a permis d'intensifier leur implication dans la réalisation des mouvements rapides et précis. Dans la présente étude, nous posons l'hypothèse selon laquelle le modèle interne de planification motrice serait modifié préalablement à l'initiation de la séquence pour tenir compte de la variabilité terminale de la plateforme de réception. Ainsi, dans une séquence de plusieurs mouvements, le phénomène de *planification couplée* sera mis en évidence et révélera une modification de la cinématique dès le premier sous-élément moteur (la saisie). De plus, cette planification couplée se traduira par une augmentation significative du lien entre les caractéristiques des sous-éléments d'une même séquence (i.e. une corrélation accrue, par exemple, entre les pics vitesse des éléments 1 : saisie et 2 : déplacement). Enfin, le caractère prédictible de l'instabilité induite au cours de l'expérience devrait avoir un effet sur la qualité de la planification. Ainsi, nous testerons l'hypothèse selon laquelle dans un environnement instable, le couplage sera d'autant plus important (lien plus marqué) si la variabilité de l'environnement est prédictible.

## **Méthodologie**

Neuf sujets sains droitiers qui ne présentent aucune altération de la vision ou de la motricité ont participé à cette expérience. Les enregistrements ont été réalisés à l'aide de 4 caméras de capture de mouvement 3D (Qualisys). Quatre capteurs réfléchissants ont été fixés sur l'effecteur de saisie : l'ongle du pouce, l'ongle de l'index, et deux marqueurs sur le poignet de chaque participant.

Pendant la session expérimentale, le sujet était debout face à une table sur laquelle est disposée un réceptacle de départ (15cm de diamètre). Ce réceptacle était placé à 35 cm de la position de départ de la main du sujet. A droite de ce réceptacle de départ, un réceptacle d'arrivée de diamètre 9cm pouvait se déplacer sur un rail de 38 cm incliné à 45° (voir figure 1). La tâche du sujet était de saisir un objet cylindrique (diamètre de 6 cm) entre le pouce et l'index, de le soulever du réceptacle de départ pour le déposer précisément sur le réceptacle d'arrivée.

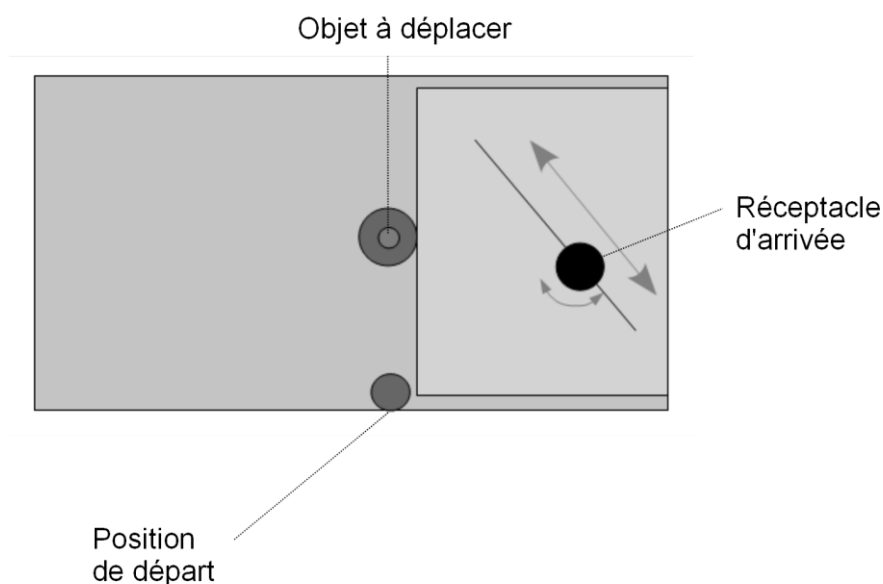


Figure 1 : Schéma du dispositif expérimental.

Deux types de contraintes ont été manipulées. Pour la contrainte vitesse, les sujets devaient réaliser la tâche à vitesse rapide (pic de vitesse du premier mouvement  $> 1200\text{mm.s}^{-1}$ ) ou à vitesse préférée. Pour la contrainte de variabilité terminale, le réceptacle d'arrivée pouvait être immobile (condition **Contrôle**) en **Translation** ou en **Rotation**. Pour la translation et la rotation, les déplacements du réceptacle pouvaient être soit prédictibles soit aléatoires.

Dans la situation prédictible, la translation consistait en un déplacement le long du rail en aller-retour avec un changement de direction toutes les secondes et une vitesse moyenne de  $50 \text{ mm.s}^{-1}$ ; la rotation consistait en un déplacement de la partie supérieure du réceptacle selon une direction comprise entre 0 et  $180^\circ$  toutes les secondes. Dans les deux cas, les conditions aléatoires consistaient en des changements non prédictibles, selon un intervalle temporel compris entre 500 et 1500ms.

Les sujets réalisaient la tâche de manipulation en situation d'interaction augmentée, i.e., sans intervention verbale de l'expérimentateur (pour une meilleure validité expérimentale). Trois petites diodes de couleur étaient placées sur la table face aux sujets pour leur indiquer le début et la fin de chaque essai. Lorsque la lumière était jaune, les sujets devaient se placer en position de départ. Après un délai aléatoire de 1 à 3 secondes, la lumière jaune s'éteignait et simultanément la lumière verte s'allumait. Lorsque la lumière verte était allumée, les sujets réalisaient l'action de prendre le cylindre et de le placer sur son réceptacle. Après 4 secondes, la lumière verte s'éteignait et la lumière rouge s'allumait. Ceci indiquait la durée inter essai au cours de laquelle les sujets devaient repositionner l'objet face à eux et pouvaient ensuite se reposer. Au bout de 4 secondes, la lumière jaune réapparaissait et les sujets recommençaient une nouvelle fois la procédure. A la fin de chaque essai, un signal auditif était émis pour indiquer au sujet si l'objet avait été correctement positionné sur le réceptacle. Les sujets réalisaient un total de 15 essais pour chacune des conditions. Pour des raisons de fatigabilité des sujets, nous avons réalisé cette expérience en deux sessions contrebalancées avec une session en vitesse rapide et une session en vitesse préférée.

L'erreur terminale a été calculée à partir des matrices de covariances des positions terminales en deux dimensions. Celles-ci étaient visualisées sous forme d'ellipses caractérisées par deux vecteurs représentant les composantes principales en longueur et en largeur. L'erreur terminale reportée ici est l'amplitude du plus grand des deux vecteurs. Pour des raisons techniques, cette erreur terminale n'a pu être calculée que lorsque le réceptacle était en condition **Contrôle** et **Rotation**. Les données cinématiques ont été enregistrées



(fréquence d'échantillonnage de 200 Hz) et analysées de façon à extraire les paramètres moteurs pour chacune des deux phases de la séquence (P1 : phase d'atteinte et saisie; P2 : phase de déplacement et pose). Ainsi, l'erreur terminale, le temps de réaction, la durée des mouvements, et la vitesse maximale ont été soumis à deux ANOVA à mesures répétées dont le facteur principal était **Stabilité** (Contrôle ; Translation<sub>Prédictible</sub> ; Translation<sub>Aléatoire</sub> ; Rotation<sub>Prédictible</sub> ; Rotation<sub>Aléatoire</sub>). Les données issues des sessions en vitesse rapide et préférée ont été traitées séparément. Le seuil de significativité a été placé à 0,05. Dans une seconde analyse, les paramètres cinématiques de pics de vitesse et de durée de mouvements ont été soumis à une série d'analyses de corrélations. Pour chaque sujet, le paramètre cinématique de la phase 1 a été corrélé avec le même paramètre de la phase 2, sur un même essai. Les coefficients de corrélation R et de détermination R<sup>2</sup> ont été calculés pour chaque condition et par sujet.

## Résultats

### Variables globales (Temps de Réaction et Erreur terminale)

En vitesse rapide, nos résultats montrent que le temps de réaction n'est pas affecté par la variabilité terminale du réceptacle que ce soit en Translation,  $F(2,6)=2.386$ ;  $p=.173$ , ou en Rotation  $F(2,6)=1.403$ ;  $p=.316$ . Le temps de réaction moyen est de  $217\pm 21$ ms en condition Contrôle,  $228\pm 22$ ms en Translation Prédictible,  $195\pm 20$ ms en Translation Aléatoire,  $241\pm 26$ ms en Rotation Prédictible et  $216\pm 32$ ms en Rotation Aléatoire. L'erreur terminale n'est pas non plus affectée par la stabilité terminale du réceptacle,  $F(2,6)=2.8185$ ,  $p=.137$ . Elle est de  $6.3\pm 0.5$ mm en condition Contrôle,  $7.0\pm 0.5$ mm en Rotation Prédictible et  $7.7\pm 0.6$ mm en Rotation Aléatoire. En vitesse préférée, des résultats similaires ont été obtenus. Il n'y a pas d'effet de la stabilité terminale du réceptacle sur le temps de réaction, ni en Translation  $F(2,16)=0.579$ ;  $p=.572$ , ni en Rotation  $F(2,16)=0.579$ ;  $p=.572$ . Le temps de réaction moyen est de  $247\pm 20$ ms en condition Contrôle,  $262\pm 19$ ms en Translation Prédictible,  $282\pm 24$ ms en Translation Aléatoire,  $281\pm 22$ ms en Rotation Prédictible et  $268\pm 20$ ms en

Rotation Aléatoire. L'erreur terminale n'est pas non plus affectée par la stabilité terminale du réceptacle,  $F(2,16)=0.769$ ,  $p=.480$ . Elle est de  $5.3\pm0.5\text{mm}$  en condition Contrôle,  $5.3\pm0.6\text{mm}$  en Rotation Prédicible et  $5.5\pm0.6\text{mm}$  en Rotation Aléatoire.

### **Effet de la stabilité terminale sur la cinématique du 1<sup>er</sup> mouvement.**

En vitesse rapide, un effet significatif de la stabilité terminale du réceptacle a été obtenu sur la cinématique du premier mouvement. Les séquences motrices étaient caractérisées par un pic de vitesse plus élevé pour un objet à placer sur un réceptacle immobile que sur un réceptacle en mouvement. Ces résultats ont été obtenus pour la Translation  $F(2,6)=31,089$ ;  $p<.001$ ,  $\eta^2_p=.91$  et également, pour la Rotation,  $F(2,6)=15,427$ ;  $p=.005$ ,  $\eta^2_p=.72$ . Pour la Translation, l'analyse post-hoc a révélé une différence significative entre la condition Contrôle et les deux types d'instabilité (Contrôle =  $1835\pm59\text{mm.s}^{-1}$  vs Prédicible =  $1543\pm65\text{mm.s}^{-1}$ ,  $p<.001$ ; Contrôle vs Aléatoire =  $1544\pm81\text{mm.s}^{-1}$ ,  $p<.001$ ). En revanche, les deux types d'instabilité ne sont pas différentes entre elles (Prédicible vs Aléatoire,  $p=.999$ ). Pour la Rotation, une différence significative est observée entre la condition Contrôle et les deux conditions variables (Contrôle =  $1835\pm59\text{mm.s}^{-1}$  vs Prédicible =  $1451\pm39\text{mm.s}^{-1}$ ,  $p=.009$ ; Contrôle vs Aléatoire =  $1482\pm55\text{mm.s}^{-1}$ ,  $p=.013$ ) mais pas de différence entre les deux types d'instabilité (Prédicible vs Aléatoire,  $p=.932$ ).

De façon similaire, les résultats ont montré que la durée du premier mouvement était réduite lorsque les sujets devaient prendre l'objet à poser sur un réceptacle immobile plutôt que sur un réceptacle en mouvement, de Translation  $F(2,6)=20,309$ ;  $p=.002$ ,  $\eta^2_p=.87$  et également, en mouvement de Rotation,  $F(2,6)=13,885$ ;  $p=.005$ ,  $\eta^2_p=.82$ . L'analyse post-hoc a indiqué qu'en condition Contrôle, la durée du premier mouvement était réduite par rapport à toutes les autres conditions d'instabilité, (1) en Translation (Contrôle= $362\pm8\text{ms}$  vs Translation Prédicible= $439\pm13\text{ms}$ ,  $p=.002$ ; Contrôle vs Translation Aléatoire= $403\pm7\text{ms}$ ,  $p=.043$ ), et (2) en Rotation (Contrôle vs Rotation Prédicible= $441\pm17\text{ms}$ ,  $p=.016$ ; Contrôle vs Rotation Aléatoire= $451\pm11\text{ms}$ ,  $p=.009$ ). De même, il n'y a pas de

différence entre les deux types d'instabilité sur la durée du premier mouvement, entre les conditions Prédictible et Aléatoire pour la condition de Translation,  $p=.062$  et également pour la condition de Rotation,  $p=.851$ .

En vitesse préférée, nos résultats ont révélé des effets similaires avec un effet de l'instabilité terminale sur le pic de vitesse du premier mouvement. Lorsque les sujets devaient prendre un objet pour le placer sur un réceptacle immobile, les mouvements étaient plus rapides que pour des mouvements vers un réceptacle instable en Translation,  $F(2,16)=11,903$ ;  $p<.001$ ,  $\eta^2_p=.60$  et en Rotation,  $F(2,16)=13,542$ ;  $p<.001$ ,  $\eta^2_p=.63$ . L'analyse post-hoc a révélé de plus une différence entre la condition contrôle et les deux types d'instabilité à la fois pour (1) la Translation (Contrôle= $1276\pm14\text{mm.s}^{-1}$  vs Translation Prédictible= $1103\pm34\text{mm.s}^{-1}$ ,  $p<.001$ ; Contrôle vs Translation Aléatoire= $1152\pm39\text{mm.s}^{-1}$ ;  $p=.014$ ) et pour (2) la Rotation (Contrôle= $1276\pm14\text{mm.s}^{-1}$  vs Rotation Prédictible= $1098\pm34\text{mm.s}^{-1}$ ,  $p<.001$ ; Contrôle vs Aléatoire= $1128\pm39\text{mm.s}^{-1}$ ,  $p=.004$ ). Les résultats n'ont pas montré de différence entre la condition Prédictible ou Aléatoire ni en Translation,  $p=.409$ , ni en Rotation,  $p=.720$ .

Pour la durée du mouvement d'atteinte vers l'objet, les résultats montrent des différences moins marquées qu'en vitesse rapide, mais qui vont dans le même sens. La durée du premier mouvement est réduite en condition Contrôle, par rapport aux conditions de Translation  $F(2,16)=3,698$ ;  $p=.048$ ,  $\eta^2_p=.32$ , et de Rotation,  $F(2,16)=3,555$ ;  $p=.053$ ,  $\eta^2_p=.31$ . L'analyse post-hoc révèle que la condition Contrôle est différente des deux autres conditions d'instabilité, pour la Translation (Contrôle= $604\pm17\text{ms}$  vs Translation Prédictible= $668\pm25\text{ms}$ ,  $p=.044$ ; Contrôle vs Translation Aléatoire= $678\pm22\text{ms}$ ;  $p=.024$ ) et pour la Rotation (Contrôle vs Rotation Prédictible= $676\pm28\text{ms}$ ,  $p=.029$ ; Contrôle vs Rotation Aléatoire= $670\pm26\text{ms}$ ;  $p=.043$ ). La figure 2 présente les profils moyennés de vitesse, pour les 15 essais de chaque condition pour un sujet type. La figure 3 présente les moyennes de groupe pour le pic de vitesse du premier mouvement.

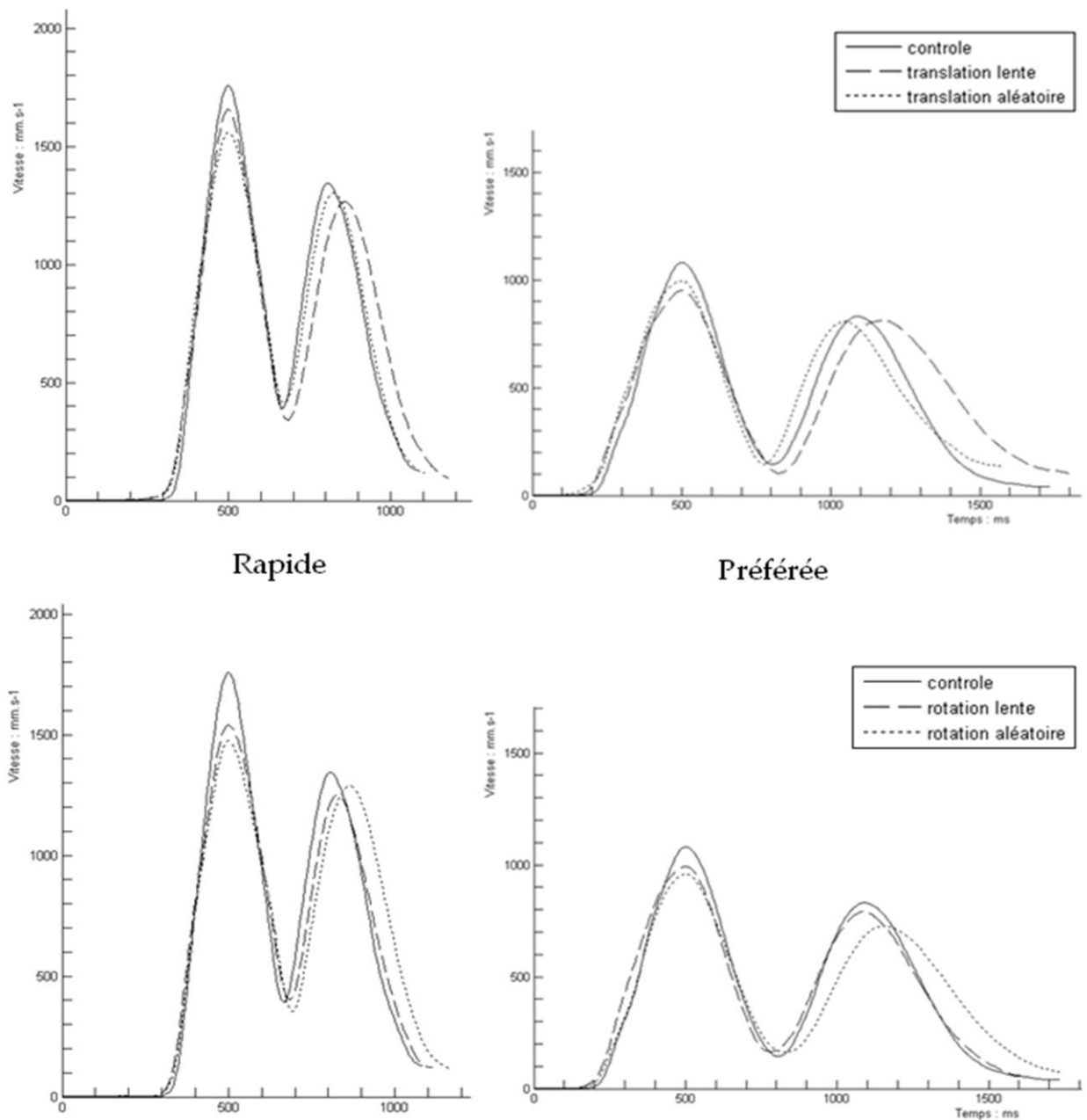


Figure 2 : Les profils cinématiques présentent la vitesse d'exécution des deux mouvements (prendre puis déposer l'objet) au cours du temps. Chacun des deux mouvements se présente sous la forme d'une courbe en cloche (phase d'accélération et de décélération). Ici la séquence des deux mouvements successifs est présentée selon les différentes conditions expérimentales.

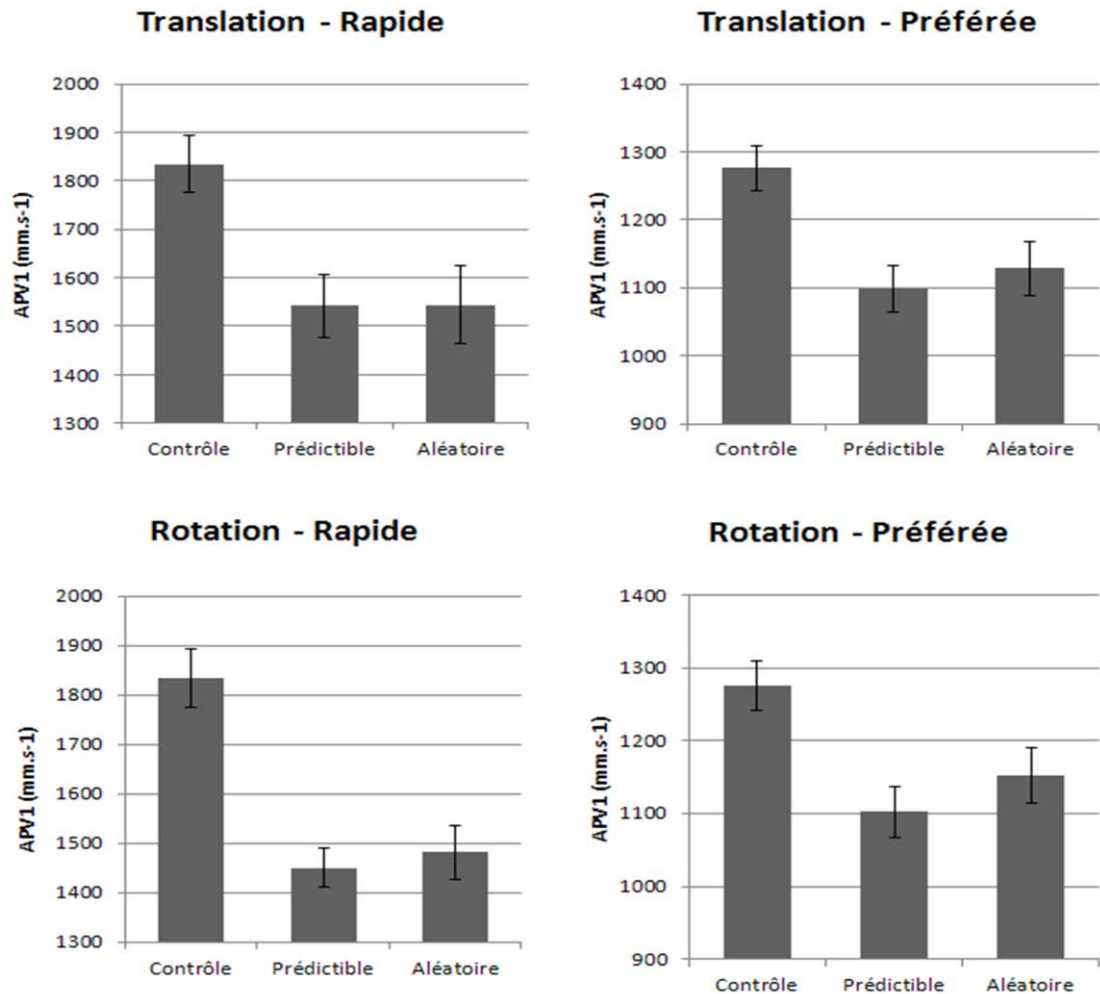


Figure 3 : Effet de la stabilité terminale sur le pic de vitesse du premier mouvement (APV1)

### Effet de la variabilité sur la corrélation entre les deux mouvements.

Dans une deuxième analyse, nous avons réalisé une analyse par corrélations pour révéler les liens possibles entre les paramètres cinématiques des deux sous-éléments d'une même séquence. Plus spécifiquement, pour les pics de vitesse en condition Contrôle seulement 2 sujets sur 9 ( $R=.34 \pm .06$ ) présentent des régressions significatives ( $R > .50$ ,  $p < .05$ ), contre 5 sur 9 en Translation Prédictible ( $R=.62 \pm .08$ ), 3 sur 9 en Translation Aléatoire ( $R=.42 \pm .09$ ), 6 sur 9 en Rotation Prédictible ( $R=.52 \pm .09$ ) et 1 sur 9 en Rotation Aléatoire ( $R=.31 \pm .05$ ). Pour les durées de mouvement, en condition Contrôle, aucun sujet ne révèle de régression significative ( $R=.28 \pm .04$ ), contre 4 sur 9 en Translation Prédictible ( $R=.42 \pm .08$ ), 1 sur 9 en Translation

Aléatoire ( $R=.28\pm.07$ ), 5 sur 9 en Rotation Prédicible ( $R=.55\pm.08$ ) et 2 sur 9 en Rotation Aléatoire ( $R=.35\pm.08$ ).

Ainsi, les résultats de l'analyse de corrélation suggèrent un effet de l'instabilité du réceptacle sur la corrélation des pics de vitesse. Dans le cas d'une Translation, les pics de vitesse sont mieux corrélés dans le cas d'une translation Prédicible ( $R^2=.44\pm.10$ ) par rapport à la condition Contrôle ( $R^2=.15\pm.04$ ),  $p=.015$ ,  $F(2,16)=5,751$ ;  $p=.013$ ,  $\eta^2_p=.41$ . En revanche, une Translation Aléatoire, n'a entraîné qu'une faible augmentation de la corrélation ( $R^2=.25\pm.08$ ), et cette différence n'est pas significative par rapport au Contrôle,  $p=.558$ . Pour la Rotation, la corrélation entre la durée des mouvements est améliorée pour une Rotation Prédicible ( $R^2=.35\pm.08$ ) par rapport au Contrôle ( $R^2=.09\pm.02$ ),  $p=.035$ ,  $F(2,16)=4,272$ ;  $p=.033$ ,  $\eta^2_p=.35$ . En revanche, une Rotation Aléatoire n'a entraîné qu'une faible augmentation de la corrélation ( $R^2=.18\pm.07$ ) qui n'est pas significativement différente du Contrôle,  $p=.606$ . En conclusion, nos résultats ont révélé une corrélation positive et significative entre les sous-éléments d'une même séquence, avec un effet amplifié (un lien plus fort) lorsque l'instabilité est prédictible. Un exemple de résultats pour un sujet type est présenté en figure 4.

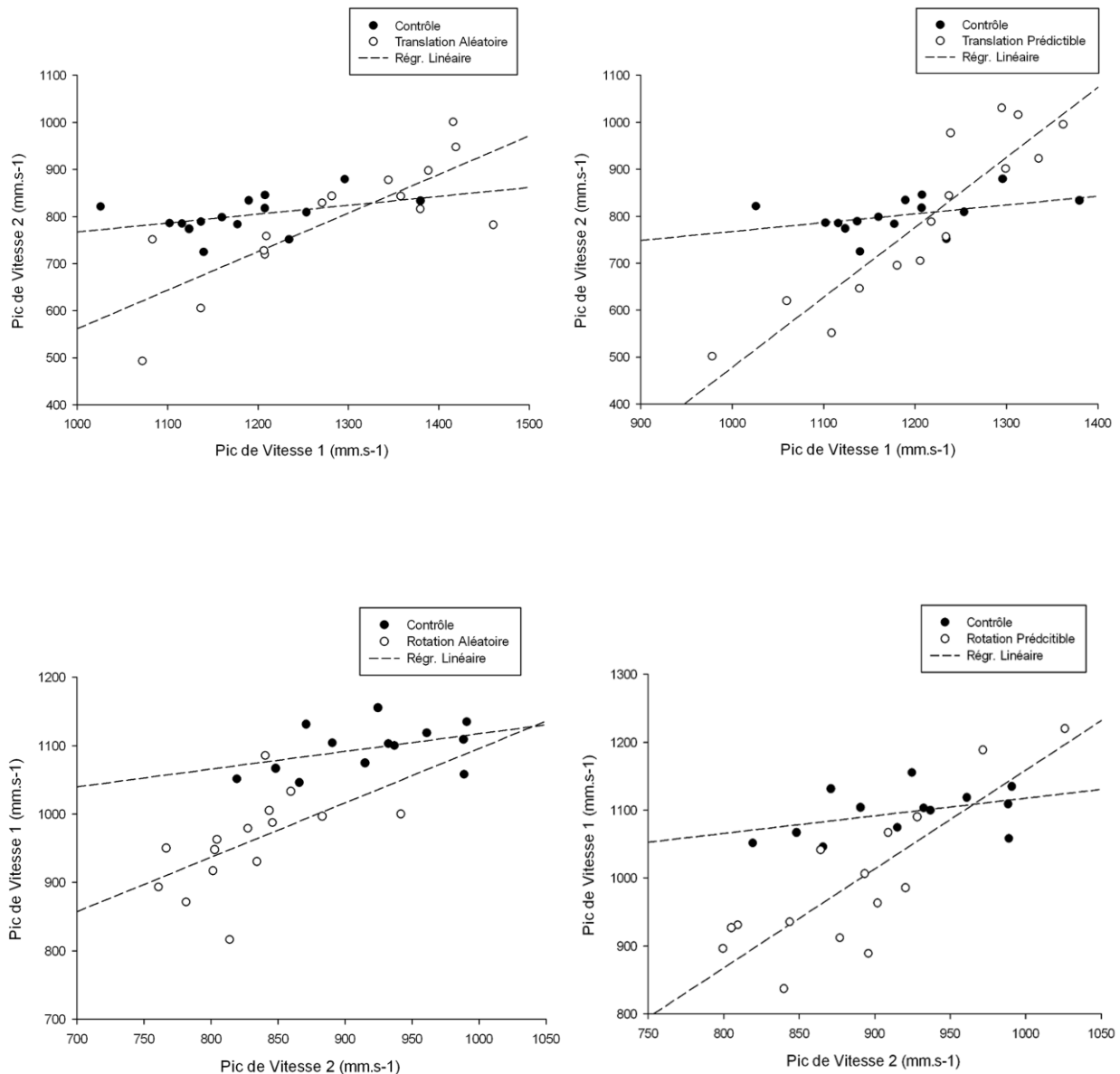


Figure 4 : Effet de la variabilité terminale sur la corrélation entre le pic de vitesse de la phase 1 et 2. En haut : corrélations en *Translation* pour le sujet JJ. A gauche : situation Contrôle en noir ( $R^2=.18$ ,  $p=.119$ ) vs Translation Aléatoire en blanc ( $R^2=.64$ ,  $p<.001$ ), à droite : Contrôle en noir vs Translation Prédicible en blanc ( $R^2=.83$ ,  $p<.001$ ). En bas : Corrélations en *Rotation* pour le sujet AR. A gauche : Contrôle en noir ( $R^2=.17$ ,  $p=.121$ ) vs Translation Aléatoire en blanc ( $R^2=.30$ ,  $p=.033$ ), à droite : Contrôle en noir vs Translation Prédicible en blanc ( $R^2=.68$ ,  $p<.001$ ).

## Discussion

Dans notre vie quotidienne, nous sommes amenés à réaliser des mouvements qui font généralement partie d'une séquence de plusieurs mouvements comme par exemple prendre une cuillère pour l'amener à la bouche ou la ranger dans un tiroir. Dans la littérature du contrôle moteur, des études ont montré qu'il est tout à fait possible de planifier une séquence d'actions selon une stratégie étape-par-étape dans l'ordre de leur exécution, mais avec une efficacité et une fluidité limitée. Ainsi, une meilleure stratégie semble être de réaliser une planification globale de la séquence motrice qui intègre non seulement l'objectif final mais également les contraintes de variabilités dont les sources sont multiples (Ansuini et al., 2006; Cohen & Rosenbaum, 2004).

Les résultats de notre étude ont permis de mettre en évidence deux phénomènes importants dans la planification d'une tâche de manipulation séquentielle impliquant deux sous-éléments. D'une part, nous avons montré que la cinématique du premier mouvement est affectée par l'instabilité terminale du réceptacle: les mouvements d'atteinte et de saisie présentent un pic de vitesse plus réduit et la durée des mouvements est augmentée si la position terminale du réceptacle est variable. Ces résultats suggèrent que l'instabilité terminale liée à l'environnement est rétro-propagée sur le premier sous-élément de la séquence afin de produire un mouvement compatible avec l'objectif moteur, i.e., placer correctement l'objet sur le réceptacle. Ces résultats sont en accord avec la littérature et confirment ceux obtenus par Rand et collaborateurs (1997) dans une tâche de pointage séquentiel montrant que les contraintes de précision terminale affectent la planification globale et l'organisation entière de la séquence (Rand et al., 1997). D'autre part, nous avons en plus démontré que l'importance de cette planification globale était d'autant plus grande que la variabilité terminale était prédictible. Ainsi, les liens de corrélation étaient significativement plus intenses entre les différents sous-éléments d'une même séquence pour des conditions de variabilité prédictible que pour des conditions de variabilité aléatoire.



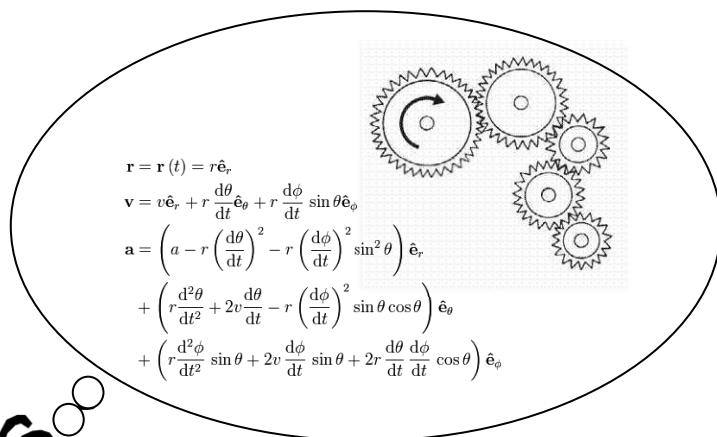
Sur le plan théorique, nos résultats ne confirment pas le modèle de planification hiérarchique (D. A Rosenbaum, Hindorff, & Munro, 1987) qui suggère que les mouvements sont planifiés dans l'ordre de leur exécution. Au contraire, ils sont en faveur du modèle de planification anticipée du mouvement (Fischer et al., 1997; Oliver Herbert & Butz, 2009) qui pose l'hypothèse selon laquelle les contraintes du second mouvement sont traitées précocement et pris en compte dès la phase de planification du premier mouvement. Par ailleurs, nos résultats semblent indiquer que les effets de variabilité liée à l'environnement sont aussi pris en compte par le système de contrôle moteur humain au même titre que les effets liés à une variabilité de localisation, de planification ou d'exécution. Dans la présente étude, nous avons proposé de manipuler la stabilité terminale de l'objectif moteur en introduisant une variabilité en translation ou en rotation. Nos résultats ont montré que dans le cas d'une translation, la variabilité *spatiale* du réceptacle est prise en compte lors de l'étape de localisation de la cible. Dans le cas de la rotation, nous avons manipulé la variabilité *temporelle* de l'instabilité du réceptacle. Dans les deux cas, les sujets ont été capables de le prendre en compte en se servant d'un mode global de planification leur permettant de réaliser la tâche sous les contraintes de rapidité et de précision imposées. Cependant, une différence significative a été mise en évidence en fonction de la prédictibilité de la variabilité. En effet, dans le cas d'une instabilité non-prédictible, nous avons montré que les sujets agissent en *réaction* face aux modifications de l'environnement et de ce fait, ne présentent pas de corrélation entre les deux sous-éléments de la séquence. En revanche, lorsque l'instabilité est prédictible, les sujets *prédisent* les conséquences de leur propre action dans l'environnement et peuvent ainsi générer très précocement des séquences motrices adaptées dans leur forme cinématique (trajectoire, vitesse, accélération). De plus, les caractéristiques cinématiques des sous-éléments sont liées entre elles pour compenser l'accumulation des sources d'instabilité et optimiser l'atteinte de l'objectif moteur.

Il est intéressant de noter que les autres études utilisant des protocoles de manipulation d'objet n'ont pas réussi à mettre en évidence une programmation couplée (Johnson-Frey et al.,

2004; Mason, 2007). La précision terminale de la séquence a été manipulée et les résultats n'ont révélé aucun effet précoce sur la cinématique du premier mouvement. Ainsi, il est possible que dans ces études, les contraintes liées à l'environnement aient été insuffisantes pour engendrer une modification précoce au sein de la planification motrice. En effet, dans les études de pointage (Rand et al., 1997), la tâche des sujets est souvent de réaliser des mouvements rapides avec une précision terminale bien contrôlée via ordinateur. Dans les études de manipulation d'objet (Mason, 2007), des instructions verbales sont utilisées, les essais sont réalisés en vitesse préférée et peu de contrôle est mis en place pour estimer la précision terminale réelle. Dans la présente étude, nous avons choisi d'utiliser un paradigme d'enregistrement, interactif pour le sujet. Un programme (Matlab) d'analyse cinématique en temps réel a permis par exemple, de moduler le retour auditif en fin d'essai en fonction de la performance du sujet. Grâce à ce système d'interaction, nous avons pu améliorer l'agréabilité de la passation expérimentale et ainsi augmenter le niveau d'implication des sujets. De façon plus importante, le système de mesure en temps réel, nous a aussi permis de vérifier la vitesse d'exécution des mouvements au même titre que les contraintes terminales pour induire une vraie situation d'exécution rapide. Dans une étude précédente, nous avons pu montrer que les contraintes de précision terminales étaient rétro-propagées uniquement en condition d'exécution rapide avec une absence de rétro propagation en vitesse préférée (Lewkowicz et Delevoeye-Turell, soumis). Nous avons alors suggéré qu'en vitesse préférée, les sujets n'utilisaient pas de stratégie globale pour planifier leurs mouvements et que l'absence de contrainte de vitesse leur permettait de planifier chaque geste séparément selon une stratégie pas à pas – qui serait probablement moins coûteuse en ressources attentionnelles. Dans la présente étude, nos résultats suggèrent que, même en vitesse préférée, il est possible de faire émerger un phénomène de planification couplée sur une séquence de manipulation d'objet si les contraintes terminales sont fortement modifiées par une variabilité spatiale et/ou temporelle du réceptacle cible. En particulier, lors d'une instabilité prédictible, les sujets vont

ralentir leurs actions de manière coordonnée pour minimiser la variabilité d'exécution de leur séquence et ainsi optimiser la probabilité d'atteindre précisément le réceptacle en mouvement.

En conclusion, grâce à l'utilisation d'un paradigme interactif et en temps réel, nous avons pu montrer qu'afin de minimiser les conséquences délétères d'une instabilité de l'environnement sur l'exécution des actions, le système de contrôle moteur humain peut intégrer les modifications environnementales lors de l'étape de planification motrice, préalablement à l'exécution. Ainsi les conséquences de cette adaptation interne sont visibles dès les premières étapes de la séquence motrice en particulier lorsque qu'une prédiction de cette instabilité est possible. Le *couplage* entre les sous-éléments d'une même séquence aurait alors deux conséquences importantes. D'une part, pour l'agent, cela lui permet de mieux anticiper les conséquences de ses propres actions au cours de leur exécution et de pouvoir s'adapter à une modification imprévisible de l'environnement via une correction entre le but prédit par sa propre exécution et le but à atteindre. D'autre part, pour autrui, cela pourrait permettre à un observateur de lire l'objectif final de la séquence motrice à partir des premières étapes de son exécution et ainsi, anticiper et interagir avec la personne qui réalise les mouvements. Ces deux capacités sont cruciales pour permettre à deux personnes d'interagir ensemble. Dans le domaine de la recherche, une meilleure compréhension de ce mécanisme de planification couplée permettra le développement de robots humanoïdes qui aurait la capacité d'interagir ou de travailler conjointement avec des personnes. Il sera possible notamment d'implémenter dans les robots, les capacités de reconnaître les mouvements humains et d'anticiper leurs actions à partir d'indices cinématiques externes non verbaux et précoces ou encore de produire des séquences de mouvements biologiques qui pourront être reconnues et lues par des observateurs humains.



$$\begin{aligned} \mathbf{r} &= r(t) = r \hat{\mathbf{e}}_r \\ \mathbf{v} &= v \hat{\mathbf{e}}_r + r \frac{d\theta}{dt} \hat{\mathbf{e}}_\theta + r \frac{d\phi}{dt} \sin \theta \hat{\mathbf{e}}_\phi \\ \mathbf{a} &= \left( a - r \left( \frac{d\theta}{dt} \right)^2 - r \left( \frac{d\phi}{dt} \right)^2 \sin^2 \theta \right) \hat{\mathbf{e}}_r \\ &\quad + \left( r \frac{d^2\theta}{dt^2} + 2v \frac{d\theta}{dt} - r \left( \frac{d\phi}{dt} \right)^2 \sin \theta \cos \theta \right) \hat{\mathbf{e}}_\theta \\ &\quad + \left( r \frac{d^2\phi}{dt^2} \sin \theta + 2v \frac{d\phi}{dt} \sin \theta + 2r \frac{d\theta}{dt} \frac{d\phi}{dt} \cos \theta \right) \hat{\mathbf{e}}_\phi \end{aligned}$$

# Chapitre 4. Observation et Modélisation

# 4.1. Reading Motor Intention through Mental Imagery.

## Abstract

Motor imagery is defined as a dynamic state during which the representation of a given motor act is internally rehearsed without overt motor output. Some evidence in experimental psychology has suggested that imagery ability is crucial for the correct understanding of social intention. The present study was conducted first to confirm that the nature of the motor intention leads to early modulations of movement kinematics. Secondly, we tested whether humans use imagery to read an agent's intention when observing the very first element of a complex action sequence. Results revealed early variations in movement kinematics and further showed that human agents can use these deviants to distinguish above chance level between three different social actions. Response times in the observation task were similar in duration to those measured in the true production task, suggesting the use of motor imagery for trial categorization. Nevertheless, in a third study, the use of an artificial (Neural Network) classifier demonstrated that classification within the first 500ms is possible without cognitive imagery processing. Hence, our results suggest that low-level motor indices afford intention reading without need for motor imagery but that human agents may use imaging beyond simulation to create an embodied sense of interactivity.

## Introduction

Motor imagery is a cognitive process in which a subject imagines that he/she is producing a movement without actually performing it and without even tensing the muscles. It is a dynamic state during which the representation of a specific motor action is internally activated without any motor output (M. Jeannerod, 1994; Lotze & Cohen, 2006). A fast-growing number of studies are indicating that brain areas engaged in the actual performance of movements are also active during motor imagery (Dechent, Merboldt, & Frahm, 2004; Gerardin et al., 2000; Grezes & Decety, 2000; Hallett, Fieldman, Cohen, Sadato, & Pascual-Leone, 1994; Hanakawa et al., 2003; Marc Jeannerod, 2001; Kimberley et al., 2006; Lotze et al., 1999; Sirigu, Cohen, Duhamel, & Pillon, 1995; Stephan et al., 1995). Besides the overlap in neural activation between imagery and execution, there are also similarities in the behavioral domain. For instance, the time to complete an imagined movement is similar to that needed for actual execution of that movement. This phenomenon known as mental isochrony has been reported in various tasks, e.g., handwriting (Decety & Michel, 1989; Decety, 1993), object manipulation (Frak, Paulignan, & Jeannerod, 2001), tapping on targets of different dimensions (Sirigu et al., 1995) and even whole body actions like walking through doorways of different widths (Decety & Jeannerod, 1995).

These findings have led to a theoretical position termed the simulation hypothesis (Marc Jeannerod, 2001) that states that movement execution and motor imagery are driven by the same basic cognitive mechanism, with motor imagery being conceived as an „offline” operation of the motor areas in the brain. Through mental simulation, motor imagery would not only produce internal images of our body in action from a 1st person perspective but it would also contain elements of kinesthetic sensations, which act as a substitute for the sensory feedback that would arise if the movement were executed (Naito et al., 2002). Hence, motor imagery is a cognitive activity that requires and uses mental simulation with in addition an *embodied* aspect, which is created through the voluntary activation of those sensory-motor areas in the brain that would be activated if we were the true actor of a movement (Mulder, 2007).

What would then be the phenomena used during the observation of biological motion? Brass et al. (2000) conducted a study in which subjects were instructed to perform a finger movement as rapidly as possible. Results indicated that the initiation times of the movements were faster when the to-be-performed action was identical to the movement that was observed, suggesting a priming effect of similarity between executed and observed actions (Brass et al., 2000). More recent studies have confirmed this effect for hand posture (Urgesi, Candidi, Fabbro, Romani, & Aglioti, 2006) and sequence learning (Pascual-Leone, 2001; Zijdwind, Toering, Bessem, van der Laan, & Diercks, 2003). More importantly for the purpose of the present study, observing an action primes the very muscles needed to perform the same action (Craighero, Bello, Fadiga, & Rizzolatti, 2002; Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Fadiga et al., 1995). Hence, observed actions activate in the observer's brain the same mechanisms that would be activated, were that action intended or imagined by the observer (Gallese & Goldman, 1998).

The question asked in the present contribution is the nature of the information that is used during action observation to permit intention reading. Indeed, chronometric studies have shown that viewers decide if the presented stimulus is a left or right hand, by engaging implicit motor imagery, a simulation that reflects in fact the time constraints of limb rotation (Parsons, 1994, 2001). It has also been shown that humans use implicit imagery when asked to decide how they would grasp a handle appearing in a variety of orientations with the subject's choices being in streaked agreement with the grips chosen during actual grasping (Johnson, 2000). However, in these studies the exact nature of the information used to simulate, imagine and then decide is not defined or discussed. In the present study, we will consider the possibility that body kinematics are characterized by early deviants that reflect the social intention of the agent. We will further demonstrate that these early deviants can be used as none-contextualized cues to predict the motor intention of an observed actor through the use of motor imagery.

Optimal control models of biological movement have been more successful than any other class of models to predict empirical findings as diverse as movement corrections from unexpected changes or responses to global perturbations (Shadmehr & Mussa-Ivaldi, 1994), but also modeling the structure of motor variability (Gordon et al., 1994; Messier & Kalaska, 1999; van Beers et al., 2004) and of generic motor laws (Lacquaniti et al., 1983). Optimal control theory helps to understand why humans move in a specific way by showing that the movements are optimal with respect to an ecologically valid criterion. But many experimental cases have shown that this optimal control is not a rule and that in many cases, naturalistic movements do not respect the optimal control laws. Some have named these exceptions as task-irrelevant factors (Oliver Herbert & Butz, 2011). A much-studied example is that of the end-state comfort effect (e.g. (D. A. Rosenbaum et al., 2012)). In this situation, the initial discomfort of the arm posture is tolerated for the sake of final comfort, because the end-state comfort is critical for good execution of future task demands (David A. Rosenbaum & Jorgensen, 1992; David A. Rosenbaum et al., 1990; Short & Cauraugh, 1997).

In the case of complex sequential actions (as reaching to use an object), it seems legitimate to ask whether the constraints of subsequent elements would have effect on the previous elements embedded in the same sequence. Thus, in an optimal point of view, the optimization criteria would not be based on performing correctly each sub-element of the sequence but rather, it would be to follow a global optimal criteria that will thus constrain the entire motor sequence and its performance. In a pick and place task, Lewkowicz & Delevoye-Turrell (under revision) manipulated the constraints set upon the final target of a two-sequence action. They reported that both first and second sub-elements of the sequence were affected by specific end-position constraints (distance, accuracy, stability). Moreover, they found that *coupled planning* (i.e. embedding the two movements together) occurred only if subjects were able to anticipate the consequences of both sub-elements prior to initiation of sequence execution. Thus, it was suggested that despite the fact that the reach-to-grasp sub-element possesses its own specific constraints (object size, relative distance and timing), the



criteria set for the second sub-element (required speed, accuracy and stability) would constitute a global and dominant speed-accuracy trade-off property that back-propagates to modulate very early on the execution of the first sub-element of the sequence.

Motor intention may also be a parameter that modifies early sequence kinematics. For example, Ansuini & coll. (2008) measured the prior-to-contact grasping kinematics for reach-to-grasp movements performed toward a bottle filled with water. By comparing hand shaping across tasks involving different subsequent actions - pour the water into a container; throw the bottle; move the bottle from one spatial location to another - the authors demonstrated how the prior intention in grasping the object strongly affected the positioning of the fingers during the reaching and the contact phase of the action (Ansuini et al., 2008, 2006). In another series of studies, Becchio et al. investigated the effects of social context on reach-to-grasp actions. They found initial adjustments reflecting specific planning strategies (Becchio et al., 2008a) as well as online adjustments (Sartori, Becchio, Bulgheroni, & Castiello, 2009) when performing under social context (Becchio et al., 2010).

Hence, there is growing evidence that both end-point constraints and social contexts affect movement kinematics early on. Recently, it was suggested that these deviants are meaningful and may be used to read motor intention. For example, when observing actions performed under social context (or not), Castiello and collaborators demonstrated that humans can successfully use kinematic cues of reach-to-grasp movements to predict the final goal of the action (Sartori et al., 2011). However, other contextual cues may have been used, e.g., the face, body movements and gaze orientation to infer the goal of the action and its social underlying intention. To study this aspect, Stapel and collaborators (2012) investigated specifically the contribution of contextual cues and found that both sources were important: intention reading was more accurate if the observed actions were placed within a meaningful context (Stapel, Hunnius, & Bekkering, 2012). Nevertheless, humans can exploit subtle movement cues alone as suggested by an elegant study using point-light displays of simple

reach to grasp movements (Manera et al., 2011). However, in this later study, choices were to be made between voluntary actions that were performed under different speeded conditions (cooperation vs. competition; fast vs. slow), which may have induced artificial kinematic differences.

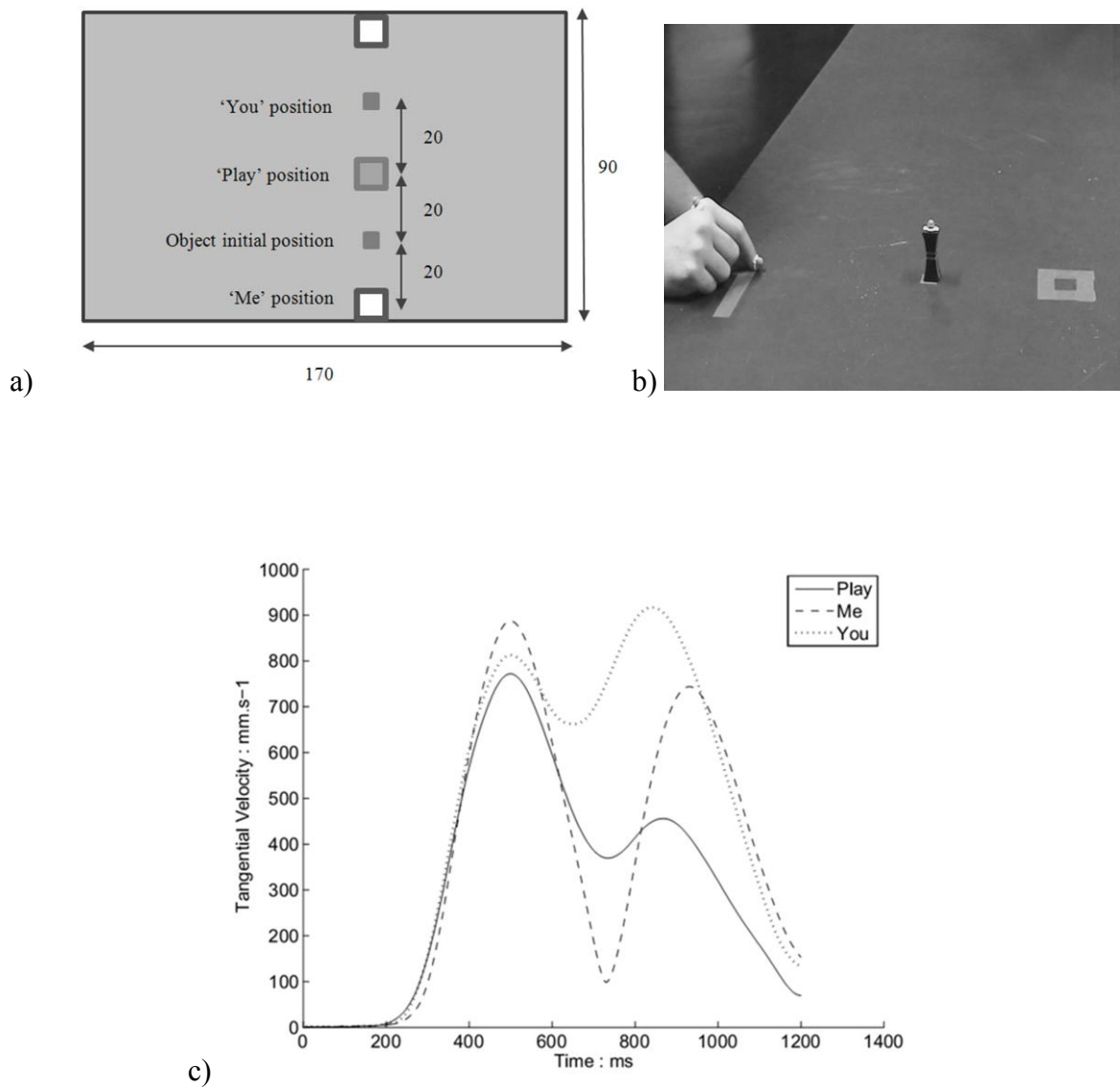
In the present study, we wanted to examine the capacity of humans to read motor intention (1) during voluntary motor sequences performed under constant speeded constraints, and to discuss (2) the need of motor imagery in the process. Complexity was manipulated both on the fact that the motor sequence was constituted of 2 elements (elt1: reach to grasp; elt2: lift to place) and was performed during a true social-interactive game with a con specific. As such, we recorded sequential actions during an ecologically inspired task (*Jungle Speed*) in a face-to-face game using a unique manipulated object. One important criterion in the design construction was to select a first common action (i.e. the reach movement) that leads to different subsequent situations, which had direct impact on the game's progression.

Confronting Jacob & Jeannerod's (2005) *reading motor intention* hypothesis, we hypothesized that human agents are able to read motor intention through the simple observation of arm kinematics of the first element of a 2-sequence action. This is possible due to the fact that arm kinematics of the reach to grasp movements reveal specific deviants in function of goal intention from an ideal optimized trajectory. Finally, if motor imagery is not necessary for intention reading, and if motor simulation is sufficient, then an artificial neural network should be able to learn from the deviants and predict as well as humans, the motor intention of an observed agent. In the following section, we first describe the methods we used to make the observation videos (Part A), which were then played to human agents (Part B) and used as input parameters to an artificial neural network (Part C).

## **Methods**

### **Part A: Creating The Stimuli**

Two adults participated in the study, one as the experimenter and the other as the subject. Both participants were right handed as verified with the Handedness questionnaire (Oldfield, 1971). They had no prior knowledge of the experiment and provided informed consent before participating in the experimental session that lasted approximately 90 minutes. The subjects' movements only were recorded using (1) a video camera (Sony Handycam) and (2) 4 Oqus infrared cameras (Qualisys system). To provide the means to analyze arm kinematics, infrared reflective markers were placed on the index (base and tip), the thumb (tip), the wrist (scaphoid and pisiform) of the subject, as well as on the object. Care was taken as to provide no contextual information within the video clips (torso, gaze, face expression), i.e., only the hand and the target object were fully in view. Cameras were calibrated before each session, allowing the system to reach a standard deviation smaller than 0.2 mm, with a 200 Hz sampling rate. Three different positions were indicated on the table-top by black tape and symbolized three specific locations that will be referred to in the next section as the placing positions: „Play', „Me', „You' (see *Figure 1a*).



**Figure 1.** **a)** Experimental Setup showing the “Play”, “Me” and “You” final position and their respective distance from the initial object position. The white squared areas are the starting hand position for both the subject (bottom) and the experimenter (top). **b)** Example of Stimuli. This close view was used for the movie clips to avoid any contextual effect (no body, no head). **c)** Typical velocity profiles observed in the „Play’, „Me’ and „You’ conditions are presented with full, dashed and dotted lines, respectively. Note that total movement time, the magnitude and the time to peak velocity of the first element of the sequence are significantly affected by motor intention (i.e., experimental conditions).

**The game.** Both the subject and the experimenter were seated at a table, facing each other. The starting position for both participants was a point placed at midline 80cm in front of the body. The object that was to be manipulated was a wooden dowel (width 2 cm ; height 4 cm) that was placed precisely 20 cm in front of the starting position of the subject. The subject's

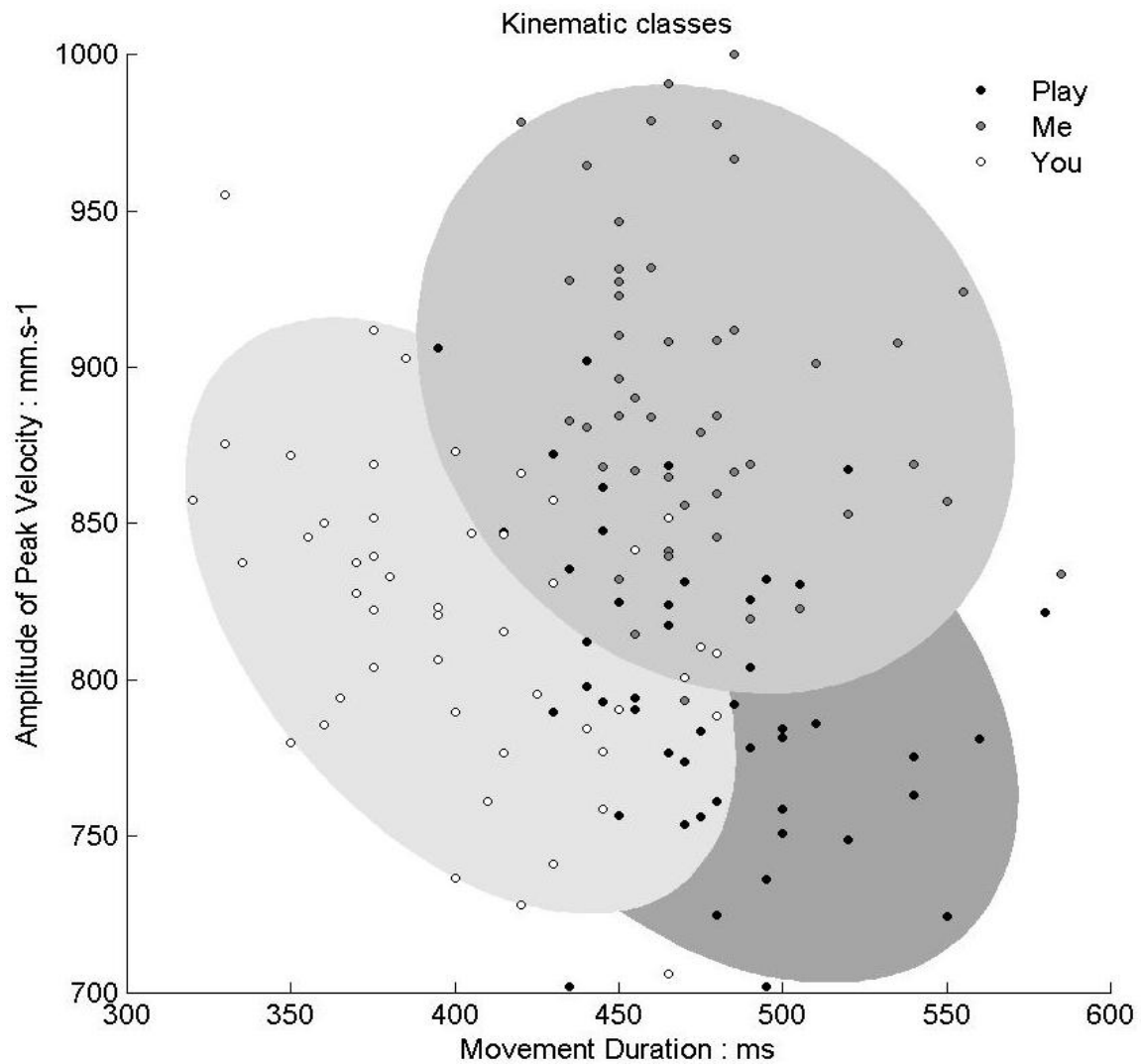
task was to reach and grasp the dowel between thumb and index finger in order to move it from the initial position to one of three placing positions, during an adapted version of the *jungle-speed* game (Asmodee eds.). A trial was defined as a series of three moves: initiate, compete, reward. The initiate-move required subjects to pick and place the dowel on the „Play’ position. Then, at a „go’ signal (given by the computer), both participants were required to perform the compete-move, i.e., to reach for the dowel as quickly as possible. For each win, 1 point was scored. The reward-move was performed by the subject who picked up the dowel to place it on the „Me’ position (if he/she had won the point) or on the „You’ position (if the point was attributed to the experimenter). Each trial started with the dowel placed at the initial position and with the participants pinching index and thumb together, with the fingertips set upon the starting position (see *Figure 1a*). Time pressure was set only on the compete-move, which was not recorded. A block ended when one of the two players reached a total of 20 points. The game consisted in 4 blocks of approximately 40 trials.

**The recordings.** The best 16 video recordings of each category („Play’, „Me’ and „You’) were selected based on their mean bit rate quality (>4500 Kbps) and kept for future use as stimuli. Each sequence included a 1-second time interval before the initial movement onset, and was cut exactly one frame before the index finger contacted the object. Movies were compressed with FFdshow codec (MJPEG) at 50 frames per second with a screen resolution of 720x576 pixels (see *Figure 1b*). Video clips were synchronized to the recordings of arm kinematics.

**Analyzing arm kinematics.** Positional data points were filtered using a dual fourth-order Butterworth low-pass filter ( $f_c = 15$  Hz; forward and backward) and tangential 3D instantaneous velocities were calculated. A threshold of 20 mm.s<sup>-1</sup> was used to determine the onset of movement. All velocity trajectories were bell shaped and consisted in two „bells’, the first corresponding to the *reach to grasp* element, the second being the *lift to place* element (see *Figure 1c*). The amplitude of peak velocity of the first element (APV1) was extracted using the local maxima (first zero-crossing of acceleration). The end of the first element was determined as the time of occurrence of the local minima (second zero-crossing of

acceleration) between the first and the second element-peaks. The duration of the first element (MT1) was calculated as the time interval between the onset and the end of the first element. Each parameter was submitted to a repeated-measure ANOVA with *Block* and *Category* as within factors; an alpha level of significance was set to 0.05.

Results revealed an absence of Block effect on APV1,  $F(2,30) = 3.056$ ,  $p = .062$ ,  $\eta^2_p = .17$  and MT1,  $F(2, 30) = 2.727$ ,  $p = .082$ ,  $\eta^2_p = .15$ , suggesting similar movement properties across block repetitions, both for amplitude of peak velocity (Block 1 :  $M = 838$ ,  $SD = 31 \text{ mm.s}^{-1}$ ; Block 2 :  $M = 824$ ,  $SD = 26 \text{ mm.s}^{-1}$ ; Block 3 :  $M = 847$ ,  $SD = 36 \text{ mm.s}^{-1}$ ) and for movement duration (Block 1 :  $M = 440$ ,  $SD = 35 \text{ ms}$ ; Block 2 :  $M = 459$ ,  $SD = 28 \text{ ms}$ ; Block 3 :  $M = 462$ ,  $SD = 24 \text{ ms}$ ). More importantly, results showed a global effect of Category both on APV1,  $F(2,30) = 58.463$ ,  $p < .001$ ,  $\eta^2_p = .80$ , and on MT1,  $F(2,30) = 55.821$ ,  $p < .001$ ,  $\eta^2_p = .79$  indicating that when subjects reached for the dowel, peak velocities were lower for the „Play’ ( $M = 796$ ,  $SD = 27 \text{ mm.s}^{-1}$ ) than for the „Me’ category. Furthermore, movement durations were longer for the „Play’ ( $M = 481$ ,  $SD = 34 \text{ ms}$ ) than for the „You’ category. The reach to grasp movement in the „Me’ category showed both higher peak velocities ( $M = 893$ ,  $SD = 37 \text{ mm.s}^{-1}$ ) and longer movement durations ( $M = 479$ ,  $SD = 22 \text{ ms}$ ) than that observed in the „You’ category (APV1:  $M = 820$ ,  $SD = 28 \text{ mm.s}^{-1}$ , MT1:  $M = 402$ ,  $SD = 19 \text{ ms}$ ). As seen in *Figure 2*, these results lead to an overlap of 30% across conditions only. Hence, through the use of these two parameters alone, a classification system should be able to categories in a predictive matter these three motor sequences, which possess distinct motor intentions.



**Figure 2.** presents the spatial representations of the first element of the sequence the 144 stimuli that were recorded using the Qualysis 3D motion capture system. The X-axis codes movement duration. The Y-Axis codes maximum amplitudes of peak velocity of the first element of the sequence. In the three „Play’, „Me’ and „You’ categories, the 95% confidence ellipses are plotted in dark, grey and white colors, respectively. The overlapping areas were estimated as following: 3 categories = 6,8%, 2 categories = 24,6%, No overlap = 68,6%.

## Part B: Human Prediction Of Ongoing Actions.

In the present study, the short video clips were presented to a panel of human subjects to test whether human agents are able to predict the goal of a sequential action when shown only the first element of a sequence, i.e., the reach to grasp element. To test the hypothesis that a

simple artificial classifier could also learn to discriminate between social categories (because the biological kinematics are affected by the agent's intention), the 3D instantaneous velocities of the arm kinematics corresponding to each short video clip were fed as input parameters to a classical feedforward neural network with one hidden layer. These experiments are presented in the following sections, *Part B* and *Part C*, respectively.

**Participants.** Twenty-six young adults (mean age:  $21.82 \pm 2.76$  years, range = 18 - 29 years) participated in the study. All subjects were right handed (Oldfield, 1971) and had no prior knowledge of the experimental goals. Subjects provided informed consent before participating in the experimental session that lasted approximately 45 minutes.

**Apparatus and Software.** Participants were seated comfortably facing a table in a dark and silent room. For each trial, participants started by placing their hand on response keys that were delimited by tape placed directly on the number-pad (numbers 2, 5, 8 were used). Stimuli were presented on a laptop computer with MATLAB software (Mathworks) in the *PsychToolbox* environment. Analogical scales (10cm lines) were used for self-evaluation of performance levels.

**Experimental Procedure.** The participants' task was to answer after each video clip presentation whether the social intention of the sequence was „**let's Play**’ (key5), „**for Me**’ (key2) or „**for You**’ (key8). A 1-second blank screen was displayed in between two trials. Participants were instructed to give their answers as fast and as accurately as possible. They were obliged to provide an answer within a 4-second time window otherwise the trial was cancelled and presented at the end of the block. A feedback message was given when responses were too slow. Each block consisted in the random presentation of a series of 48 stimuli, i.e., 16 different video clips for each of the three categories (Play; Me; You). At the end of the block, a 5-minute pause was systematically taken. At this occasion, participants filled in an analogical scale to provide a subjective judgment of the performance they thought to have achieved on a scale from 0 (not at all) to 10 (absolutely).

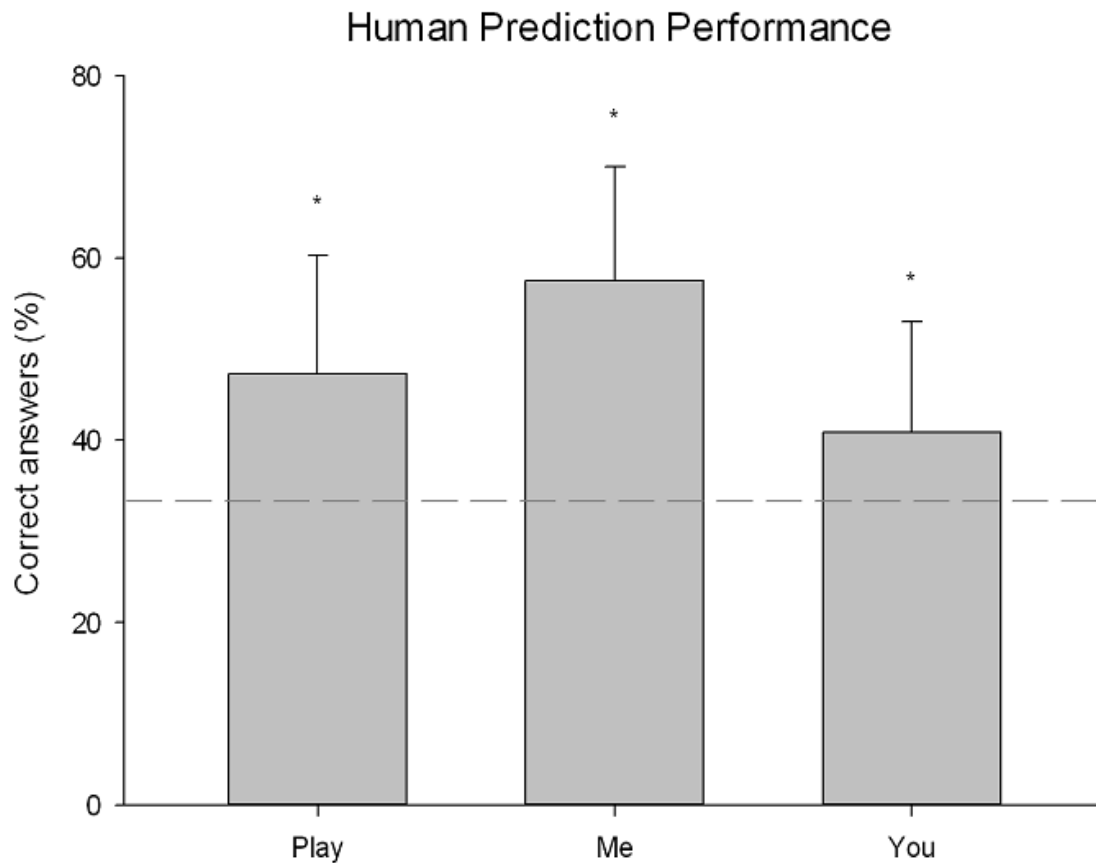


**Dependent variables and statistical analyses.** For each trial and participant, response times were calculated as the time interval between the presentation of the last frame of the video and the subjects' key press. Mean percentages of correct responses, mean response times and mean self-evaluation scores were calculated for each category and submitted to a repeated-measure ANOVA with *Block* and *Category* (Play; Me; You) as within factors. We also conducted sub-analyses. (1) For the percentages of correct responses, scores for each category were compared to the reference constant, i.e., the random answer value of 0.33, with a single sample t-test. (2) To gain an indicator of motor imagery, response times were compared to the true movement duration presented on the video (MT1+MT2) for each category. In all of these analyses, the alpha level of significance was set to 0.05.

**Response times.** Statistical analyses revealed an absence of Block effect,  $F(2,50) = 1.401$ ,  $p = .256$ , indicating that participants answered as fast in Block 1 ( $M = 878$ ,  $SD = 382$  ms), in Block 2 ( $M = 848$ ,  $SD = 315$  ms), and in Block 3 ( $M = 944$ ,  $SD = 316$  ms). An absence of main effect was also found for Category,  $F(2,50) = 2.621$ ,  $p = .083$ , indicating that participants answered as fast for „Play’ ( $M = 900$ ,  $SD = 294$  ms), „Me’ ( $M = 866$ ,  $SD = 294$  ms), and „You’ categories ( $M = 905$ ,  $SD = 300$  ms). The sub-analysis revealed an absence of difference between response times and movement times for both „Play’ ( $M = 900$  vs  $898$  ms),  $t(25)=0.25$ ,  $p=.980$ , „Me’ ( $M = 866$  vs  $964$  ms),  $t(25)=1.694$ ,  $p=.103$  and „You’ categories ( $M = 905$  vs  $907$  ms),  $t(25)=0.032$   $p=.975$ . These findings indicate that the participants waited a time interval before giving their response that corresponded to the duration of true execution of the entire motor sequence (reach, graps, lift and place) suggest the use of motor imagery for decision making.

**Percentages of correct responses.** There was an absence of Block effect on classification performances,  $F(2,50) = 0.102$ ,  $p = .903$ . However, a main effect of Category was obtained,  $F(2,50) = 16.022$ ,  $p < .001$ ,  $\eta^2_p=.39$ . Post hoc Scheffé analyses further showed that participants were more accurate for trials in the „Me’ category ( $M = 57.53$ ,  $SD = 13.02$  %) than in the „You’ ( $M = 40.87$ ,  $SD = 12.12$  %) and in the „Play’ category ( $M = 47.27$ ,  $SD =$

13.04 %). Single sample tests confirmed that each category was significantly greater than random chance (‘Me’:  $t(25) = 5.463$ ,  $p < .001$ ; ‘Play’:  $t(25) = 9.914$ ,  $p < .001$ ; ‘You’:  $t(25) = 3.169$ ,  $p = .004$ ). These results are presented in *Figure 3*.



**Figure 3.** Mean percentage of correct responses (standard deviations as error bars) obtained in the experiment reported in part B of the manuscript. The dotted lines represent the random baseline level of 33,33%. To note: \* $\alpha=0.05$ ; \*\* $\alpha=0.01$

**Self-evaluation.** The main effect of Block did not reach significance for the self-evaluation scores,  $F(2,50) = 2.674$ ,  $p = .079$ . Across blocks, participants self-evaluated their own performances as only slightly better than chance ( $M = 40.90$ ,  $SD = 13.15$  %). As correct classification was obtained for all categories, statistical analysis confirmed that participants significantly underestimated their true performance level,  $F(1,25)=11.680$ ,  $p=.002$ ,  $\eta^2_p=.32$ ;

this was especially true in Block 3 in which performance level was subjectively reported as being the worst ( $M = 38.17$ ,  $SD = 19.30\%$  vs.  $M = 48.80$ ,  $SD = 12.46\%$ ).

## **Part C: Predicting Sequential Actions Using Artificial Neural Networks**

In the following section, we describe the simple feedforward neural network that was developed in the present study to demonstrate the possibility to anticipate the motor intention of an agent without the need of high order cognitive processes.

**Architecture and Learning procedure.** A simple classification Neural Network was constructed with  $N$  neurons (1-23 neurons) as inputs, 3 hidden neurons and 3 output neurons (one for each category). The  $N$  size is the number of samples taken from the kinematics, which may be used for the classification task, i.e., the sub-selection of the total movement duration. Activation functions for the output layers were symmetrical and sigmoid, between -1 and 1. It is the case that the output vector can be interpreted as a logical value (+1) when the associated category is activated; the logical value (-1) is attributed otherwise. In the present case, a single component was set to +1.

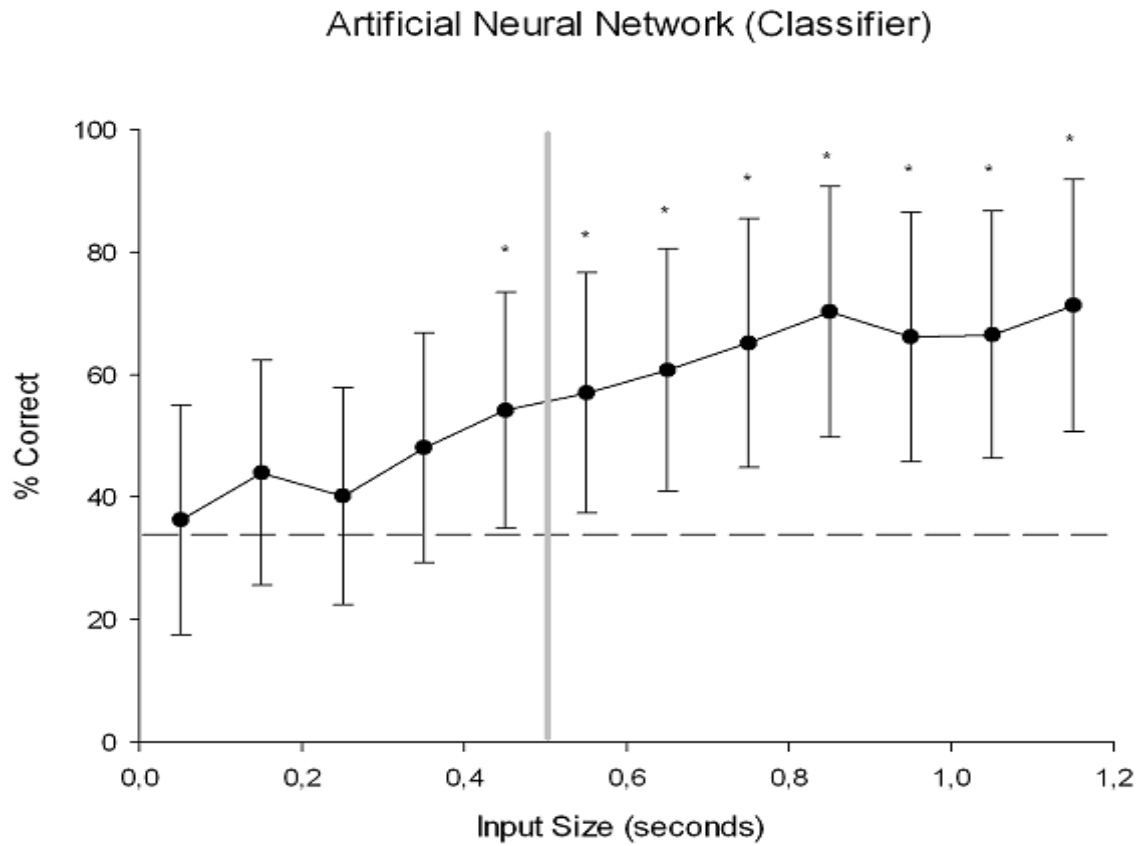
For each trial, only the section for which there was a hand-movement was considered, i.e., from the start of the first element to the end of the second element of the sequence. This time interval was determined in reference to the  $20\text{mm.s}^{-1}$  threshold criteria used in part B of the present study. The instantaneous velocity in 3D was then calculated between the two subsequent frames of the wrist's position. A *sampling* parameter was used to compute the average velocity across 10 frames. This *sampling* parameter is a constant time window of 50 ms without overlap that echoes the perceptual binding of motion events. Thus, the procedure afforded a vector of  $N$  samples that were then used as input parameters for the network. Finally, a training-set (25%) and a test-set (75%) were randomly picked from the 144 different kinematic recordings. For each possible size of the input vector (i.e. time window for kinematic recognition), 20 different networks were trained to obtain a classification

performance. The results for mean responses and variances across the 20 networks are described in the result section as the NN success rate (this value is always lower than the best performing network).

By varying the amount of data fed as input parameters (1-23), we computed the classification performance from multiple time windows (50-1150 ms). The learning procedure that was chosen is a back-propagation algorithm using the FANN library (Nissen, 2003). Target error (to stop the learning) was set to Mean Standard Error (MSE)  $< 0.001$ , with a maximum number of epochs set to 10 000, and 300 iterations between each test for the evaluation of target global error. In a preliminary study, we used networks with an augmented number of hidden units and good performances were also obtained. Nevertheless, as an ANN with 3 hidden units succeeded in learning the classification task, we opted in the present report for the simplest system in order to minimize risks of over learning.

**Classification results in function of time.** The artificial classifier was able to converge in most cases. The classifier succeeded in discriminating between categories for input sizes above 9, i.e. with at least 450 milliseconds of movement information. For the input size of 9, single sample t-tests confirmed that all categories were above chance level,  $p < .001$ : „Play’ category ( $M = 55.70$  SD = 8.08 %); „Me’ category ( $M = 56.70$  SD = 4.16 %), and „You’ category ( $M = 50.33$  SD = 5.63 %). Figure 4 presents the detailed results obtained for 12 different input sizes, between 1 (50ms) to 23 (1150ms). From the input size of 5 (250ms) to 9 (450ms), only 2 categories were successfully recognized while the other remained below chance level; Below the input size of 5, only one category was correctly classified. Depending on initial conditions and learning procedures, we observed that the categories that were more rapidly identified could change. Hence, further work is required for further interpretation of these category-specific effects. Nevertheless, the crucial point to note here is the fact that by 450ms all categories were classified above chance level; a point in time that occurred before the end of the first element of movement sequence (see Figure 4) confirming the capacity of a

simple network to categorize motor intention through the use of low-level kinematics, early on during motor execution.



**Figure 4.** Mean percentage of correct classifications (standard deviations as error bars) obtained with artificial neural networks. The horizontal axis codes the input size (step1: 50ms to step23: 1150ms). The vertical axis codes the mean and standard deviation values of the 20 networks. The dotted lines illustrate random baseline. The vertical grey bar indicates the end of the first motor element of the sequence.

## Discussion

Recent studies in robotics, cognitive sciences, and motor control have demonstrated high-skilled robotic systems that are able to solve complex tasks such as navigation, object recognition and even fine object manipulations. The development and application of e.g. the optimal theories of motor control (Doya, 2000; Todorov & Jordan, 2002; Todorov, 2004) have provided the means to develop systems that have even mastered the problem of gravity, allowing for the execution of complex physical motor sequences such as walking and

dancing. In contrast, robotic system still reveal limited adaptive capacities with humans, especially for rhythmic situations for which the *robots lack crucially of interactivity*. The absence of interactivity may be due to the fact that the optimal theories of motor control are no longer sufficient (Sisbot, Marin-Urias, Broquere, Sidobre, & Alami, 2010) to account for the behavioral data that have been reported in experimental psychology in cases of human social interactions. Indeed, an increasing number of studies are reporting that for humans acting alone (motor interaction), the motor actions performed do not follow kinematics which are similar to those used when the person acts in collaboration/competition with a partner (social interaction - (Becchio et al., 2010). These behavioral deviants that are observable in the early stages of motor execution may play a functional role and be used by con-specifics as social cues to infer motor intention. Thus, they need to be considered when creating cybernetic systems that afford true human-robot interactivity (Andry, Gaussier, Moga, Banquet, & Nadel, 2001; Gaussier, Moga, Quoy, & Banquet, 1998).

In the present contribution, we report experimental data confirming that motor intention modifies movement kinematics within the first hundreds of milliseconds. More specifically, our findings demonstrated first that the three different motor intentions that were created using a simplified version of the *Jungle Speed* game modified the kinematics of the first (reach) element of the motor sequence. Second, human agents were able to classify rapidly ( $<1s$ ) and above chance level ( $>40\%$ ), the trial category through the observation only of the reaching movement of the sequence. Response times suggested the use of motor imagery by our human participants. Nevertheless, using a classic feedforward neural network, findings also demonstrated that motor imagery is maybe not necessary since the ANN was able to categorize trials through the use of low-level kinematics within the first 450ms of the sequence. In the following section, we discuss these findings in more detail and open the discussion around the need to use motor imagery by human agents in order to experience true interactivity.

***Kinematics reflecting motor intention.*** In the abundant literature of manipulative actions, the effects of end-point constraints on the early parts of movement kinematics have been investigated extensively. In none-social situations, multiple parameters have been reported to modify and shape hand trajectory in two-element sequences such as second-target distance (Gentilucci et al., 1997), end-target orientation (Haggard, 1998; Hesse & Deubel, 2010; Seegelke, Hughes, Schütz, & Schack, 2012) and second-action type (Armbrüster & Spijkers, 2006; Johnson-Frey et al., 2004; Marteniuk et al., 1987; Mason, 2007; P. H. Weiss, Jeannerod, Paulignan, & Freund, 2000). In social tasks, final-goals have also been reported as having an effect on reach-to-grasp kinematics such as giving vs. placing an object (Becchio et al., 2008a), cooperative vs. competitive actions (Becchio, Sartori, Bulgheroni, & Castiello, 2008b; Georgiou, Becchio, Glover, & Castiello, 2007), absence vs. presence of social request (Ferri et al., 2011; Sartori et al., 2011), and even verbal communicative vs. non-communicative intentions (Sartori, Becchio, Bara, & Castiello, 2009). The kinematic effects reported in the present study are consistent with this literature and suggest that when planning a sequential action with multiple motor elements, the requirements of the endpoint element are back propagated to constrain the way the very first element of the sequence will be planned and performed. Thus, it is possible to suggest that low-level motor components may contain early indices that reflect the end-point motor intention of an agent.

***Reading motor intention...*** In the present study, each trial started with the similar movement of reaching to grasp an object, whether the action was then to give, keep or displace the object on another position on the table. Indeed, subjects initiated their move with their hand placed on the starting pad of the playing area, and reached for and grasped the wooden-peg that was always located at the same position on the table. However, the second part of the sequence was specific and directly related to motor intention: lift the wooden peg to take it („Me’ category), to give it („You’ category) or to place it on the table („Play’ category). Thus, any kinematic deviants observed on the first part of the sequence may be related to the social intention of the second part. By measuring two basic motor parameters (*peak velocity* and

*movement duration*), we showed that it was possible to dissociate the three types of social interaction categories (*Figure 2*). We then tested the fact that human observers could use these deviants to classify observed actions above chance level. The video clips were created in order to show the first element only, without any contextual cues; care was also taken to cut the end of the reaching action, one frame before object contact, in order to avoid providing any cues on movement direction of the second element of the sequence. Even if participants found the task very difficult and thought to have responded randomly, our findings demonstrated that classification is possible and that in certain cases, the participants' performance can be extremely precise (up to 67% of correct classification for the best of participants). But how is this possible?

... ***through mental imagery***. According to the simulation hypothesis (Jeannerod 2004), both observing and imagining an action activates the same neural correlates than actual execution. This possibility has since been confirmed in behavioral experiments that have shown that the time to complete an imagined movement is similar to that needed for actual execution of that same movement. When human subjects were required to classify observed actions, response times were found to be significantly longer than classic simple reaction times ( $> 500$  ms). They were in fact as long as the durations of the actual second motor element, with an absence of time differences. These findings suggest that subjects were performing motor imagery of the second motor element in order to simulate the motor intention of the observed agent, motor image on which they base their decision. It has been proposed that cognitive high-level functions like motor simulation would also be the basis for the emergence of intention understanding in more complex situations like role playing, theory of mind or empathy (Gallese, Keysers, & Rizzolatti, 2004; Gallese, 2001; Iacoboni, 2009). It would be interesting in a future study to include personality and individual characteristics to assess how emotional valence and social similarity may modulate the capacity to read intention through action observation.



***An alternative low-level hypothesis.*** It is nevertheless possible that the understanding of motor intention is based on more low-level cue readings. Indeed, despite a total absence of contextual cues within the video clips (body, head, eyes), we demonstrated in the present study that participants were able to read motor intention significantly above chance level. Hence, it might be that the subjects' responses are guided only by the slight kinematics deviances from the optimal trajectory. This would trigger a stimulus-response type of mechanism that would take place after years of interactive experiences; none cognitive by nature, it would require little resources but could lead to high recognition errors, especially in poorly known environments. As a first investigation of this alternative low-level hypothesis, we reported here a second method of investigation (presented in part C) for which we used a very simple neural network (NN) classifier and we showed that this NN was able to categorize the three categories of social interaction above chance level. The NN stabilized within the first 450 ms, which suggested that the classification was terminated before the end of the first motor element of the sequence. The performance level reached by the NN was similar to that observed in human individuals suggesting that stimulus-response coding could be sufficient for intention reading. However, it is to note that the NN was fed with extracted tangential velocity samples as input. Thus, the NN on the one hand and the humans on the other reached similar performance levels using different strategies to solve the puzzle of motor goal inference. Future studies are now required to investigate further the nature of the information that provides the best classification of motor goals/intentions. In addition, it will be important to determine whether human individuals could reach similar performance levels using direct coding of the low-level kinematic parameters (see the resonance theory by Viviani, 2002) or whether the kinematic deviants are simply a by product and hence, even for the simplest actions, humans need to engage in a cognitive simulation process to understand motor intention (for a debate see e.g. (Kilner, Friston, & Frith, 2007; Kilner & Frith, 2008).

It is to note that correct classification of the three social categories was far from being perfect, reaching in the best of cases 60% of correct identification. Hence, kinematics can be

used for predicting ongoing actions but cannot be the only source used by human agents to judge motor intention. It has been shown that during natural sequential tasks (i.e. preparing a sandwich), eye movements are stereotyped and predictive (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Pelz, Hayhoe, & Loeber, 2001), with the eyes preceding the hand movements in a systematic way (R. S. Johansson, Westling, Bäckström, & Flanagan, 2001). It is thus possible that using both gaze position and the hand kinematics, an observer is able to increase the efficiency of intention reading ((Bekkering & Neggers, 2002).

***Perspectives for interactive and social robotics.***

The application of our work would be to develop robots that afford true interaction, i.e., being able (1) to read motor intention in human kinematics in order to adapt but also (2) to move with biological realistic kinematics, in order allow others to understand the intention of the robot. The Aibo robot designed by Sony is a good example of the limitations of current approaches. The robot is quite impressive during the first minutes of functioning (smooth movements, nice physical design...) but rapidly people stop trying to interact with Aibo. Following the data presented here, we hypothesize that a humanoid robot could become interactive if it moved following the laws of biological movement with action sequences that integrate back propagation of terminal intention. Such a phenomenon would provide the means for human agents to read intentionality and thus, gain in understanding the goal of the robot's movements. Furthermore, including social deviants in the motor kinematics within early steps of motor sequences would also allow safe interaction with large industrial robots by affording humans the possibility of anticipating false moves in joint actions that share similar work spaces.

Implementing robots with the architecture necessary to “afford intentionality” would need to integrate the different brain regions that are known to play a role in motor planning and motor-sensory predictive mapping. De Rengervé et collaborators (de Rengervé, Hirel, Andry, Quoy, & Gaussier, 2011) have recently reported on such an architecture, which included amongst other areas, the cerebellum and the basal ganglia. Tested on both software and hardware, this neural architecture has demonstrated its efficiency on data collected in a

hydraulic robotic arm. With a series of imitation trials, this system demonstrated the capacity to learn how to perform sequential actions that respected biological laws, i.e., to perform movements with kinematics that mirror those performed by human agents. As such, this robot arm has demonstrated increased interactivity with human agents affording augmented interaction both in time and in space (none published results). Ongoing studies are now being conducted to assess whether this interactivity is associated to an increase in the capacity of human collaborators to read the robot's intention.

## **Conclusion**

We have reported experimental data demonstrating that it is possible to read motor intention through the simple observation of kinematic deviants. Classification capacities were significantly above chance level and provided human subjects the means to dissociate between three different socially oriented actions. We argue in the present study that reading intentionality may not depend on a high-level cognitive function as suggested in the psychological literature. Internal simulations may not be systematically required and understanding other intentions may, in certain cases, relate to a direct coding of those kinematic deviants that back propagate from end-point to early on during sequence execution. This direct coding would emerge through years of learning, during interactions with adult con-specifics. As a first step to support this hypothesis, we report in the present study simple neural networks that were able, after learning the meaning of kinematic deviants, to classify the three categories of actions to the *same degree of accuracy as the human participants, without the need of complex cognitive processes*. These preliminary results stress the importance of further development of the optimal theories of motor control to include the more cognitive aspects of social context.



## 4.2. Eye movement strategies to Read

### Intentions

#### Introduction

Trying to understand actions of others always start by observing it. The capacity to read intention of observed actions is an amazing capacity of the human brain and yet the underlying mechanisms remain to be explored. The question of how humans are able to interpret external cues of events when observing an action to form a structured representation of goals and subsets of means to achieve it is still unclear. In a previous study we have shown that it is possible to read intentions by observing only the first element of a 2-elements sequential action (Lewkowicz and Delevoye 2013). The results suggested that humans can read intentions without contextual information (body, face, etc...) based on the slight deviations from an optimal strategy in the kinematics of the same reach-to-grasp motor element but followed with three different end goals for the placing movement. Even though they were not fully aware of their capacity, the classification rates were above chance level. One possible explanation is that participants were able to automatically simulate the action in their own motor repertoire, And by the use of a predictive sensorimotor knowledge (forward model), had access to the goal of the observed action. But given the slight amplitude of the deviations from the optimal model, the inherent motor variability of the end effector, and the strong inference process we asked them in this task, according to the fact that they don't believe there were any differences in the stimuli we showed, it is the case that the subject classification rates were quite low. Nonetheless, one another possible explanation for the large number of errors would simply be that the participant were not precisely looking at the stimulus or at specific parts of the stimulus that convey the relevant information for the task.

Recent studies have shown than when observing other people interacting with the environment, humans systematically fixate action goals ahead of time. Evidences has been

found that when executing an action, we usually drive our gaze to the target long before the hand contact with it (R. S. Johansson et al., 2001; Neggers & Bekkering, 2000). This also has been found in object manipulations in natural environments (Land & Furneaux, 1997; Land & McLeod, 2000) or everyday activities (Hayhoe & Ballard, 2005; Hayhoe et al., 2003; Land, Mennie, & Rusted, 1999). This proactive gaze related to visuo-manual coordination has given even more recent interest when researchers have found that even during observation of object manipulation tasks, the observers gaze was also a systematic proactive sequential structure (Webb, Knott, & MacAskill, 2010) and tends to mimic gaze position of the agent (R. S. Johansson et al., 2001; Rotman, Troje, Johansson, & Flanagan, 2006). Because of the proactive nature of eye movement both in execution and observation, it support the direct matching hypothesis (Flanagan & Johansson, 2003). This hypothesis postulates that observing actions performed by others elicits a motor activation in the brain of the observer similar to what occur when planning his/her own actions (Rizzolatti et al., 2001; Rizzolatti & Sinigaglia, 2010). Thus, eye movement programs when observing actions are driven by the observer's own motor representation of the observed action. More specifically, Ambrosini, Costantini and Sinigaglia (2011) have shown that participants who observed an actor reaching for and grasping an object showed higher accuracy and earlier saccadic movements when participants observed an actually grasping hand than when they observed a mere reaching hand devoid of any kind of target related pre-shaping (Ambrosini, Costantini, & Sinigaglia, 2011). Their results are in agreement with the results of Rotman and al. which showed that under high uncertainty regarding the following target, people tend to adopt a default strategy, saccading first toward the most salient stimulus and then shifting to the other one as they realize the first wasn't the actual target (Rotman et al., 2006). On the contrary, Ambrosini and al. suggests that when motor cues are available and help selecting action targets, people automatically tend to use such motor information and be more fast and accurate at gazing at the object (Ambrosini et al., 2011). We then hypothesized that, in our experiment, when participants

used external kinematics cues for intention reading, the proactive gaze behavior would be more accurate and more stable.

The second question we asked in this study is the effect of the camera angle or perspective view on understanding intentions. Recent studies have found that humans have a spontaneous tendency to take the agent perspective when describing the relationship between objects (Furlanetto, Cavallo, Manera, Tversky, & Becchio, 2013; Tversky & Hard, 2009) suggesting that perspective-taking subserve understanding potential action (Tversky & Hard, 2009), or planning a complementary response (Furlanetto and al. 2013). Additionally, Anquetil and Jeannerod (2007) found that mental simulation of observed actions in first person or in third person shared the same proprieties such as an increase of response time as with an increase of task difficulty (Anquetil & Jeannerod, 2007). These results emphasize the idea that action execution, observation or simulation are represented in a coordinate system that could be easily transformed from one point of view to another. However, we didn't find any study that have extensively tested whether the perspective view has or has not an effect on action observation or understanding.

We thus focus on those two aspects by carrying out an experiment where participants had to watch small video clips of reach-to-grasp movements that were followed by three different placing action. The video clips stops when the hand contacted with the object and the task was to assess on which of the three different placing location the object would be placed. Participants watched clips from three different perspectives and we simultaneously recorded their eye movement during the movie presentation with an head-free eye tracking device.

## **Methods**

***Creating the Stimuli.*** To provide consistent stimuli that reflect an ecological interactive situation, we created a simple competitive game during which two individuals

were required to reach for a wooden dowel placed at the center of a table, as fast as they can before their opponent. However, we were specifically interested in the non-competitive part of the game such as setting up the game, or handing out rewards that were executed before or after the competitive part. This was done for at least two reasons. First, because the two opponents were not aware of the specific recording of the non-competitive parts, we were hoping that it would reveal more spontaneous sequences of interactive movements. Second, the competitive move does not have the same proprieties as compared to the non-competitive ones (traveled distance, speed constraint). Thus, care was taken that during the non-competitive parts of the game, the reach-to grasp elements were maintained similar (initial hand position, object position), while the placing elements of the sequences were of three different category: set up the game („play’), give a reward to me („me’) or give a reward to you („you’).

***The game.*** Both the actor and the experimenter were seated at a table, facing each other. The starting position for both participants was a point, placed at midline 15cm in front of the body. The object that was to be manipulated was a wooden dowel (width 2 cm ; height 4 cm) that was placed precisely 20 cm in front of the starting position of the subject. The actor’s task was to reach and grasp the dowel between thumb and index finger in order to move it from the initial position to one of three placing positions. Each round of the game consisted of a succession of three moves: initiate, compete, reward. The initiate-move required subjects to reach for the dowel and place it on the „Play’ position in order to set up the initial condition of the game. Then, at a „go’ signal (given by the computer), both participants were required to reach for the dowel as quickly as possible. For each win, 1 point was scored. The reward-move was always performed by the actor who picked up the dowel to place it on the „Me’ position (if he/she had won the point) or on the „You’ position (if the point was attributed to the experimenter). Each trial started with the dowel placed at the initial position and with the participants pinching index and thumb together, with the fingertips set upon the starting position. Three different positions were indicated on the table-top by black



tape and symbolized three specific locations that will be referred to in the next section as the placing positions: „Play’, „Me’, „You’ (see *Figure 1*). No time pressure was set on the non-competitive parts, which were recorded. A block ended when one of the two players reached a total of 20 points. The game consisted in 3 series of approximately 35 trials. The movements were recorded using (1) a video camera (Sony Handycam) and (2) 4 Oqus infrared cameras (Qualisys system). Care was taken as to provide no contextual information within the video clips (torso, gaze, face expression), i.e., only the hand and the target object were fully in view. Between each series of recordings the camera angle was modified according to the three different positions: “front”, “lateral” (right side) and “top” (*Figure 1*). To provide the means to simultaneously record arm kinematics, infrared reflective markers were placed on the index (base and tip), the thumb (tip), the wrist (scaphoid and pisiform) of the subject, as well as on the object.

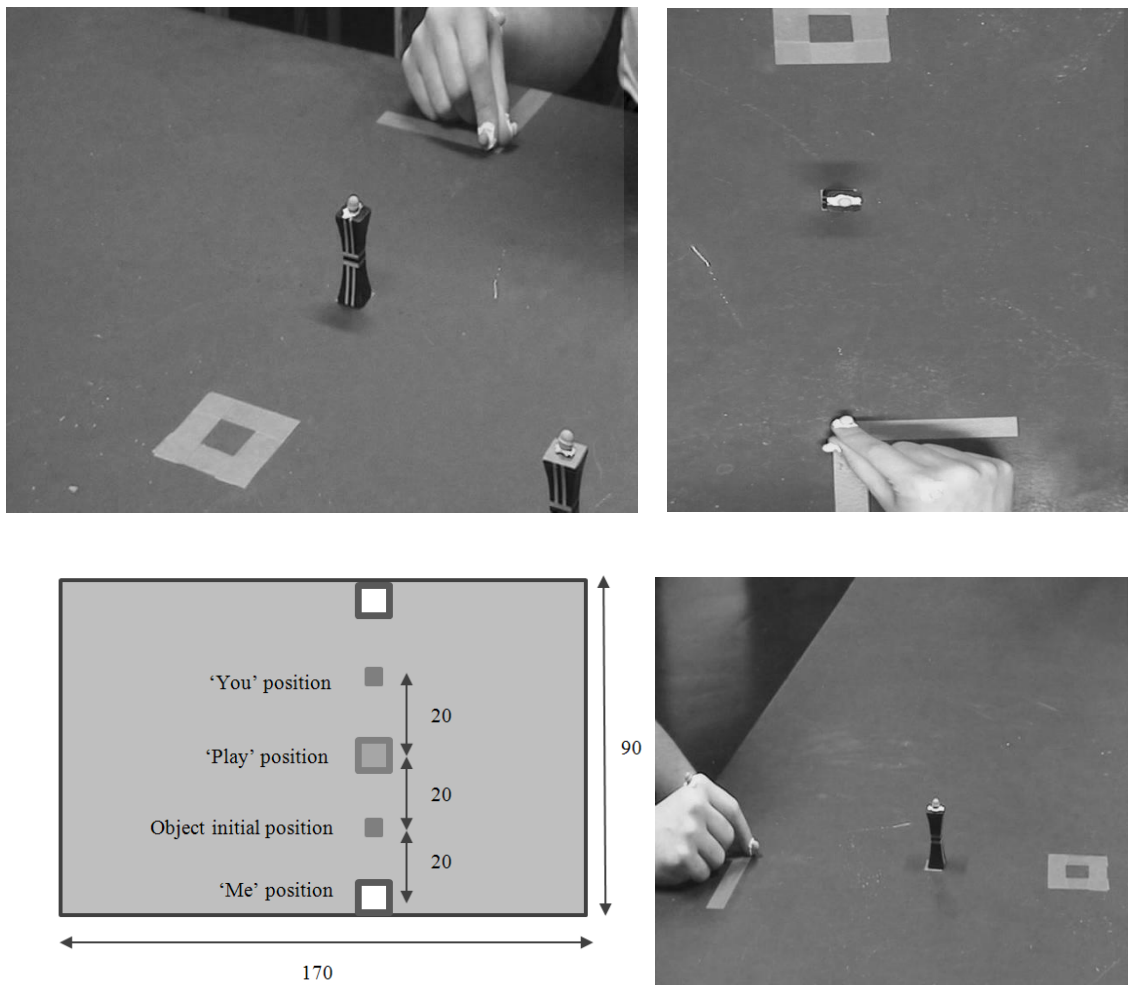


Figure 1. Examples of stimuli from different camera's angle : “front” (top-left), “top” (top-right) and “lateral” (bottom-right). Bottom-left: Experimental Setup showing the “Play”, “Me” and “You” final position and their respective distance from the initial object position. The white squared areas are the starting hand position for both the subject and the experimenter.

***The recordings.*** The best 16 video recordings of each category („Play’, „Me’ and „You’) and for each camera angle („Front’, „Lateral’ and „Top’) were selected based on their mean bit rate quality (>4500 Kbps) after video compression and kept for future use as stimuli. Each sequence included a 1-second time interval before the initial movement onset, and was cut exactly one frame before the index finger contacted the object. Movies were initially recorded at 25 fps with a screen resolution of 720x576 pixels. An deinterlace filter algorithm (virtualdub) was first applied to reach a framerate of 50fps, then the video were compressed with FFdshow in MJPEG format with an overall JPEG quality of 95%.

***3D animations.*** In order to be used as “virtual” stimuli in experiment 2 and 3, the movements were animated in 3D-space with ®MotionBuilder (Autodesk) based on kinematic recordings of the actor movements. First a 3D model of a human hand (LibHand – <http://www.libhand.org>) was imported on MotionBuilder. Then, for each recording we used the recorded position at 200 Hz (previously exported in .c3d file format) and we assigned the measured movements of the wrist, index and thumb to the corresponding part of the animated hand. The rotation movement data of the hand was applied under MotionBuilder via a “rigid body” anchor point created in respect with the two markers of the wrist and the marker at the

base of the index (meta-carpophalangeal joint). It is to note that in order to verify in a more controlled situation whether the camera angle would have an effect on the observed behavior; we used the exact same kinematics' source for both three angles. Within the 3D virtual environment, we were able to place more accurately the virtual camera at 3 different locations. We then were able to create three different animations in 2D format of the same 3D kinematics, one from each camera view. The 3D rendered animations had the same duration than their "real" equivalent movie, with a resolution of 720x576 pixels at 60fps. The latter was chosen to better match the computer screen refresh rate in order to avoid any synchronization problem.

***Participants.*** Twenty-five young adults participated in three experiments. 15 participated in Experiment 1, 25 in Experiment 2 and 10 participated in Experiment 3. All subjects were right handed (Oldfield, 1971) had no prior knowledge of the experimental goals, had normal or corrected-to-normal vision and did not report or exhibit any obvious neurological or motor deficit. The local university ethics board approved the experiments, which complied with the Declaration of Helsinki. Subjects provided informed consent before participating in the experimental session that lasted approximately 45 minutes.

***Apparatus.*** Participants were seated comfortably facing a table in a dark and silent room. For each trial, participants started by placing their hand on response keys that were delimited by tape placed directly on the number-pad (numbers 2, 5, 8 were used). Stimuli were presented on a laptop computer with MATLAB software (Mathworks) in the *PsychToolbox* environment. Analogical scales (10cm lines) were used for self-evaluation of performance levels. A RED eye-tracking device (SMI) placed under the screen was used for the recordings of gaze direction at a frequency of 500Hz. A calibration procedure was made for each participant before the experimental session allowing the system to reach an average precision of 0.5° of ocular angle.

**Procedure.** The participants' task was to answer after each video presentation whether the social intention of the sequence was „**let's Play**’ (key5), „**for Me**’ (key2) or „**for You**’ (key8). A blank screen was displayed in between two trials. After 500ms, a fixation cross appeared. Participants were required to fixate and stay within a window of 100x100 pixels around the cross for at least 500ms for the trial to begin. In experiment 2, this cross was at a constant distance from critical fixation's landing sites, i.e., the hand or the object. In experiment 3, the cross was located at the object position. This procedure ensured that in experiment 2 participants always were required to make at least one saccade in order to fixate the object or the hand. Moreover, this procedure gave us the possibility to verify that participants were correctly looking at the screen when the movie presentation began. Participants were instructed to give their answers as fast and as accurately as possible by pressing the keypad. They were obliged to provide an answer within a 4-second time window otherwise the trial was cancelled and presented at the end of the block. A feedback message was given when responses were too slow. Each block consisted in the random presentation of a series of 48 stimuli, i.e., 16 different video clips for each of the three categories (Play; Me; You). At the end of the block, a 5-minute pause was systematically taken. At this occasion, participants filled in an analogical scale to provide a subjective judgment of performance quality on a scale from 0 (not good at all) to 10 (absolutely perfect).

In experiment 3, the participants had to do this procedure twice. In one condition their eyes were free to move anywhere („free’). In the other condition, since the fixation cross was located at the object position, participants were told to maintain their eyes at the same position throughout the presentation of the movie clip. A real-time procedure was used to verify whether the eyes were correctly positioned within the 100x100 pixels window around the target. If the gaze direction was at a distance of more than 50 pixels during the movie presentation, the movie was automatically stopped and a blank screen was presented during the “missed” movie duration. During the blank screen (the time inter-trial), participants were free and even encouraged to move their eyes.

***Dependent variables and statistical analyses.*** Mean percentages of correct responses, and mean self-evaluation scores were calculated for each category and submitted to a repeated-measure ANOVA with *Block* and *Category* (Play; Me; You) as within factors. The kinematics characteristics of the arm movements that were used as stimuli here were reported in another paper (Lewkowicz et al. 2013). For the eye movements, the number of fixations and saccades events, the position and the duration of the fixations, the distance of the fixations relative to the initial hand position or relative to the object position, the onset of the saccade events relative to the hand movement as well as the time spent on the object or the initial hand position were calculated for two separate experimental time windows. More specifically, each movie clip included a 1-second duration during which the hand was not moving. We thus separated the eye-tracking measurements for *before* the hand was moving and after, i.e., *during* hand movement.

In experiment 1, we tested whether the number of saccades and fixations were the same for those trials that were correctly classified versus those trials that were not correctly classified. Thus, the previously mentioned parameters were submitted to a repeated-measures ANOVA with *Angle* („front’, „lateral’ and ‘top’) and *Response* („correct’, „incorrect’) as within factor across all subjects.

In experiment 2, we compared the classification rates and the eye-movement strategy between the “real” and the “virtual” stimuli. Thus, the parameters were submitted to an ANOVA with *Response* („correct’, ‘incorrect’) and *Type* („natural’, ‘virtual’) as within factor.

In experiment 3, we compared the effects of imposing a highly constrained strategy on subjects’ eye movements. Thus, the parameters were submitted to a repeated-measure ANOVA with *Angle* („front’, „lateral’ and ‘top’) and *Strategy* („free’ or „imposed’) as within-factors.

## **Results**

In the present study, the short video clips were presented to a panel of human subjects to test whether human agents are able to predict the goal of a sequential action when shown only the first element of a sequence, i.e., the reach-to-grasp element. We simultaneously record the eye movement during the task to test the hypothesis 1) whether the eye movements on such clips reproduce eye movements during execution 2) whether a “virtual” movie will influence the eye movements as compared to a “real” movie 3) whether constraining the eye movements will improve performance of categorization. These three experiments are presented in the following sections.

## ***Experiment 1***

### **Categorization responses**

In an overall view, the total global percentage of correct answer is above random chance level of 33% ( $M = 44.44\%$ ,  $SD = 8.81$ ),  $t(14) = 4.884$ ,  $p < .001$ . The ANOVA results showed a main effect of Angle on performances,  $F(2, 28) = 4.412$ ,  $p = .021$ ,  $\eta^2_p = .24$  indicating that classification rates were higher for the ‘lateral’ angle ( $M = 48.61\%$ ,  $SD = 11.15$ ), than for the ‘front’ angle ( $M = 44.30\%$ ,  $SD = 9.88$ ) or the ‘top’ angle ( $M = 40.42\%$ ,  $SD = 11.18$ ). A main effect of Category was also obtained,  $F(2,28) = 4.463$ ,  $p = .020$ ,  $\eta^2_p = .24$ , indicating that participant were more accurate for the ‘Me’ category ( $M = 52.78\%$ ,  $SD = 18.60$ ) than the ‘Play’ ( $M = 42.64$ ,  $SD = 6.29$ ) or ‘You’ categories ( $M = 37.91$ ,  $SD = 15.33$ ). Moreover, an interaction effect was found between the two main factors,  $F(4,56) = 2.983$ ,  $p = .026$ ,  $\eta^2_p = .18$  indicating the effects are not strictly independent. Post-Hoc Scheffé analysis indicated that there was a significant difference between the ‘Me’ and ‘You’ category when movies were presented in ‘top’ ( $p = .015$ ), and ‘lateral’ ( $p = .018$ ) angles, while this is not the case for the ‘front’ angle ( $p = .99$ ).

### **Number of Saccades.**

#### ***During Hand Movement***

Although a small number of saccades (less than a quarter of the total number of saccades) were executed during the hand movement, our results showed some interesting significant effects. First, the camera *Angle* affected the number of saccades,  $F(2,28) = 7.959$ ,  $p = .002$ ,  $\eta^2_p = .36$ , with a higher number of saccade for the ‘front’ ( $M = 32.5$ ,  $SD = 8.24$ ) and the ‘lateral’ ( $M = 30.33$ ,  $SD = 11.11$ ) views than the ‘top’ view ( $M = 20.63$ ,  $SD = 12.26$ ). More importantly, we found a global effect of the *Response* which indicates that correct responses trials have a systematic lower number of saccades during hand movements ( $M = 24.42$ ,  $SD = 9.04$ ) than the incorrect responses trials ( $M = 31.22$ ,  $SD = 10.67$ ),  $F(1,14) = 5.083$ ,  $p = .041$ ,  $\eta^2_p = .27$ , and regardless of the view angle. No interaction effect was found between *Angle* and *Response* effects,  $F(2,28) = 1.229$ ,  $p = .308$ .

### *Before Hand Movement*

Interestingly, the same pattern of results was found for the number of saccades during the 1-second time of video before the hand was moving. In overall, because the fixation cross was located away from the two landmarks, at least one saccade had to be executed to look at one or the other landmark. Thus, more saccades were executed during this period than during the hand movement. More specifically, the *Angle* showed a significant effect,  $F(2,28) = 15.634$ ,  $p < .001$ ,  $\eta^2_p = .53$ , indicating that the total number of saccades was higher for the ‘front’ ( $M = 79.23$ ,  $SD = 10.73$ ) and the ‘lateral’ ( $M = 75.70$ ,  $SD = 9.28$ ) than the ‘top’ angle ( $M = 64.30$ ,  $SD = 11.75$ ). But more importantly, the *Response* showed a significant effect,  $F(1,14) = 6.121$ ,  $p = .027$ ,  $\eta^2_p = .30$  with the ‘correct’ response trials presenting less saccades ( $M = 65.18$ ,  $SD = 14.39$ ) than the ‘incorrect’ trials ( $M = 80.98$ ,  $SD = 15.79$ ). An interaction effect was found between *Angle* and *Response* effects,  $F(2,28) = 4.198$ ,  $p = .025$ ,  $\eta^2_p = .23$ . The post-hoc Scheffé analysis indicated that the *Response* effect was most present on the ‘top’ *Angle* (correct:  $M = 50.07$ ,  $SD = 17.09$  vs incorrect:  $M = 78.53$ ,  $SD = 21.83$ ),  $p = .006$ . The two others Angles showed an increased number of saccades during incorrect trials, but the increase were not significant (Figure 2).

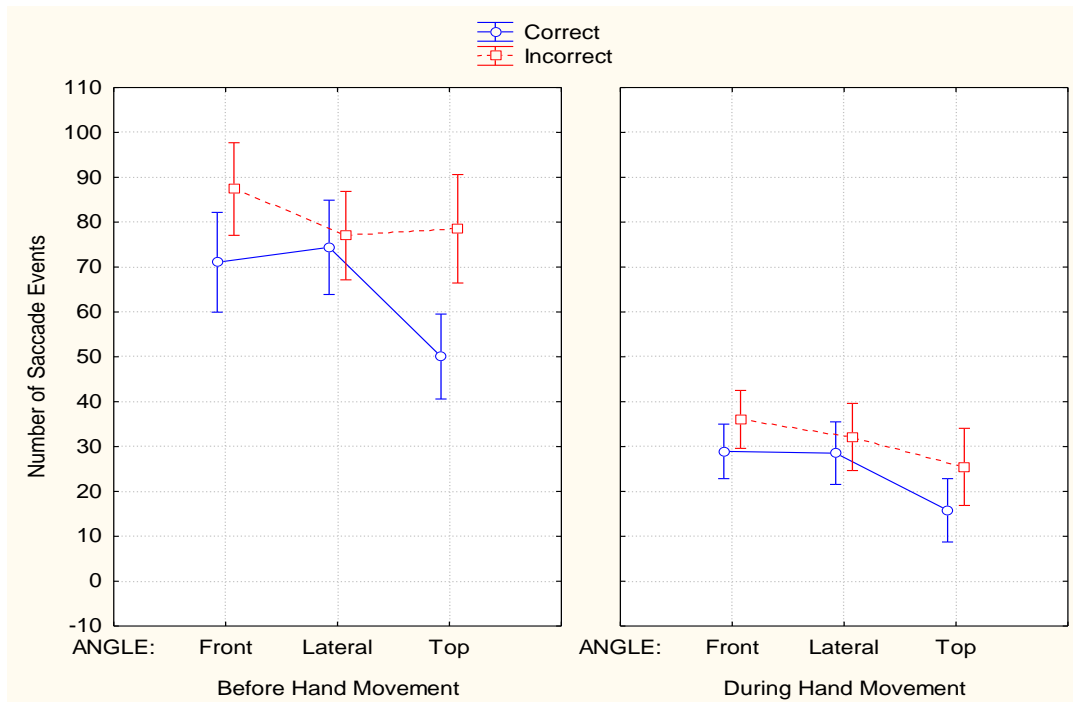


Figure 2. Number of Saccade Events

Taken together these results showed that the total number of saccades events is strongly influenced by the perspective showed to participants, but also that these events occurrence will influence the final classification rates. At first, it seems that the more they will move their eyes, the less their classification rate will be and this can be seen both *during* and *before* the hand movement. However, because we believe this is an incomplete description, we need to go further into the results to better understand the eye movement behavior on this “anticipation” of task.

### Location of Fixations and Dwell times.

In order to better understand what specific content in the movies would drive the observer gaze we separate the fixations that landed at the vicinity of the object that was to be grasped, from the ones that landed at the vicinity of the hand initial position. We measured the Euclidian distance between the fixation and the reference point and we used a simple criteria of 100px around the target for the fixation to be “at the vicinity” of the target or not.

### *Dwell times during Hand Movement*



During the hand movement, we only report fixation that landed at the vicinity of the object for two reasons. First, the „landmark’ of the hand would be changing in position and size for every frame of the movie, it would be difficult to have a precise measure of the distance between the actual fixation and this landmark. Second, because this „landmark’ would be moving, the eyes won’t really make „fixations’ if they look at it, but rather some „smooth pursuit’ which are not easy to detect precisely. We thus focus on fixations on object and we reported on the following parts the fixations at the vicinity of the hand initial position only for the time *before* the hand movement. The mean duration of fixations on the landmark: „object’ are affected by the camera’s *Angle*,  $F(2,26) = 20.661$ ,  $p < .001$ ,  $\eta^2_p = .61$ . Post-hoc Scheffé analysis indicate that the ‘top’ *Angle* showed a longer mean durations for fixations ( $M = 910\text{ms}$ ,  $SD = 265$ ) than the „front’ ( $M = 486\text{ms}$ ,  $SD = 131$ ),  $p < .001$  or the „lateral’ *Angle* ( $M = 602\text{ms}$ ,  $SD = 176$ ),  $p < .001$ . Neither effect of *Response*,  $F(1,13) = 1.700$ ,  $p = .215$  nor any interaction effects between experimental factors,  $F(2,26) = 1.062$ ,  $p = .360$ , were found for this measurement.

#### *Dwell times before Hand Movement*

Before the hand is moving, our results showed the same pattern of results, with an effect of *Angle*,  $F(2,26) = 18.075$ ,  $p < .001$ ,  $\eta^2_p = .58$ . On overall, the dwell times on the object *before* the hand movement are shorter than *during* hand movement but the „top’ view still have higher dwell times ( $M = 235\text{ms}$ ,  $SD = 27$ ) than „front’ ( $M = 203\text{ms}$ ,  $SD = 31$ ),  $p < .001$  or „lateral’ ( $M = 189\text{ ms}$ ,  $SD = 29$ ),  $p = .002$ . The dwell times at the vicinity of the hand showed no effects of *Angle*,  $F(2,20) = 0.501$ ,  $p = .613$ , or *Response*,  $F(1,10) = 0.064$ ,  $p = .805$  (Figure 3)

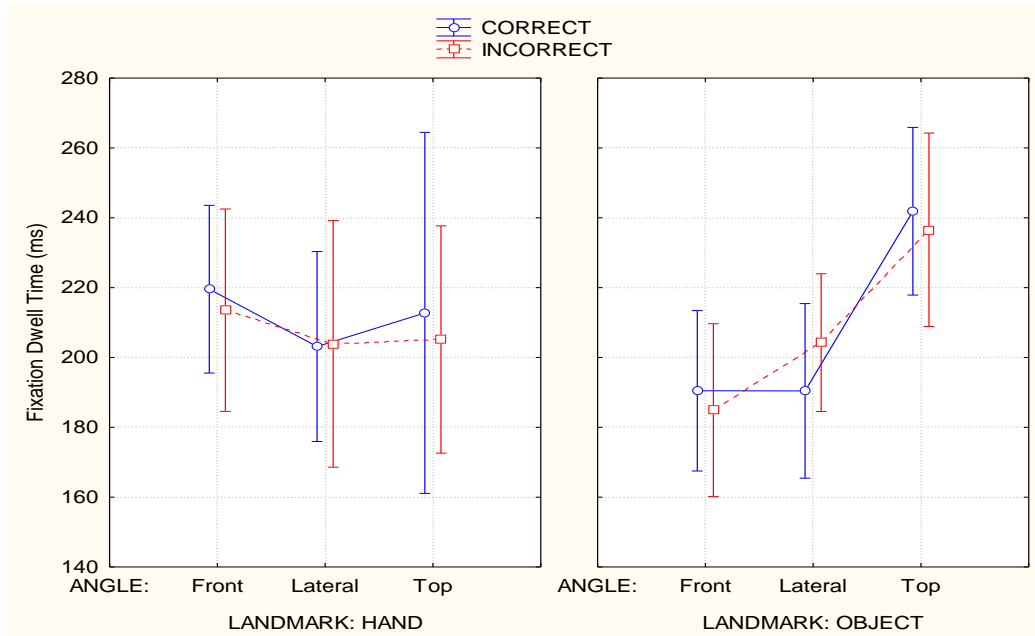


Figure 3. Dwell Times of fixations on the Hand or the Object

These results indicated that the dwell time for all fixations is globally the same, regardless if the response was correct or incorrect. Thus the time spent on the object or on the hand is not sufficient to predict if the movement will be correctly anticipated. In other words, to give a correct answer, a participant doesn't need to look at the object or the hand for a longer time. However, the perspective given by the camera *Angle* will strongly influence the dwell times, indicating that the eye movements were also influenced by the low-level statistics of the movies.

### Saccade Onset.

As reported by earlier studies (Johansson et al. 2003, Rotman et al. 2006), most of the saccades were executed before the hand movement in an „anticipative' fashion. Nonetheless, across all subject, a little less than one quarter of the saccades were initiated after the initiation of hand movement, and we showed that these „late' saccades executed in a „reactive' fashion would give rise to a larger proportion of wrong categorization response. To go a little further, we specifically look at the onset time of the saccade that will conduct to the fixation event on the object that is maintain during the hand movement. And, we separated the saccade executed before the hand movement with the saccades executed during the hand movement.

### *Saccade Onset during Hand Movement*

For those saccades that were not executed before the hand movement, we found no effect of *Angle*,  $F(2,22) = 0.011$ , no effect of *Response*,  $F(1,11) = .656$ ,  $p = .435$  or any interaction between the two experimental factors  $F(2,22) = 2.118$ ,  $p = .144$ .

On overall, these results indicate that the onset of saccade during the hand movement ( $M = 259\text{ms}$ ,  $SD = 47$ ) is made regardless of the *Angle*, or the *Response* („correct’ or „incorrect’). As mentioned earlier, the occurrence of “late” saccades is higher for incorrect trials; however we showed here that these “reactive” saccades had the same overall latency on correct or incorrect trials.

### *Saccade Onset before Hand Movement*

Our results showed a significant effect of *Angle* on the saccade onset,  $F(2,24) = 19.990$ ,  $p < .001$ ,  $\eta^2_p = .62$ , indicating that the saccades occurred later in the „Front’ view ( $M = 666\text{ms}$ ,  $SD = 126$ ) than in the „Lateral’ ( $M = 587\text{ms}$ ,  $SD = 154$ ) or the „Top’ view ( $M = 409\text{ms}$ ,  $SD = 136$ ). No effect of *Response*,  $F(1,12) = 0.749$ ,  $p = .404$  or interaction between *Angle* or *Response*,  $F(2,24) = 0.442$ ,  $p = .648$  were found (Figure 4).

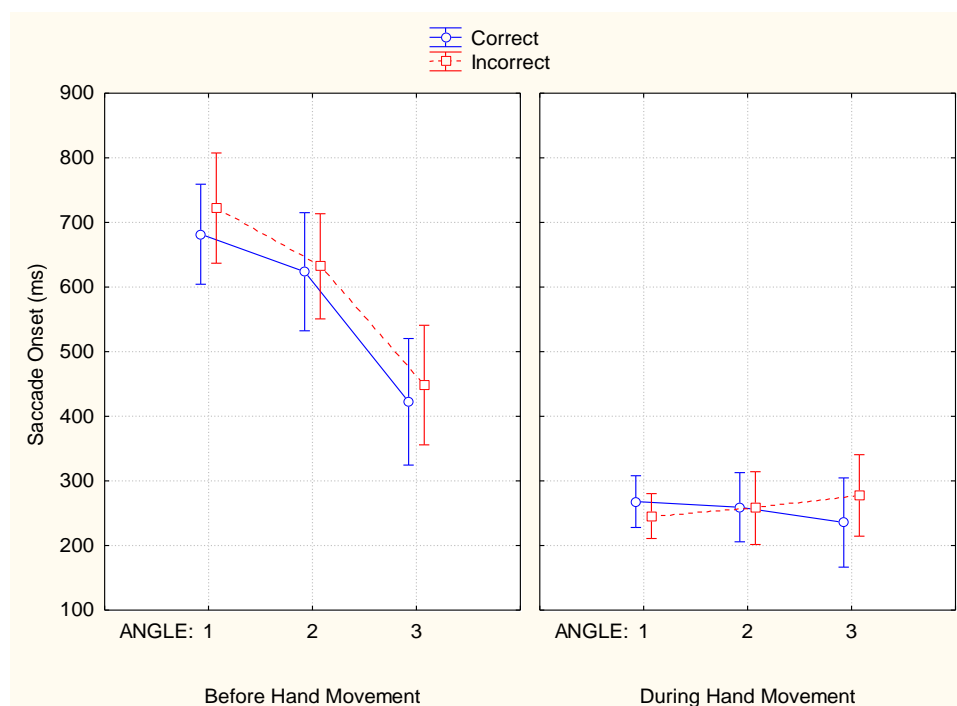


Figure 4. Saccade Onset

Here we found no *Response* effect, indicating that the onset of the “anticipative” saccade was not earlier in correct trials than in incorrect trials. This could go against the basic assumption that “the more I look, the more I get”, but it doesn’t. Actually, we reported an effect of *Angle* indicating that the „Top’ view showed earlier saccades. We also reported that the „Top’ view had the lowest classification rates (40.42%), suggesting that the quality of information retrieved by the subject is different regarding the stimulus perspective. Thus, the early onset of the saccades could be due to this lower overall information quality of the stimulus rather than a specific strategy regarding to the anticipative nature of the task. To summarize, if the timing proprieties of eye movements are not related to the classification rates, we then tested if the spatial proprieties would have been predictive of a correct or an incorrect answer. We then tested the spatial variability of the fixations at the vicinity of the object.

### **Fixation Dispersion.**

We assessed the scattering of fixation’s positions at the vicinity of the object during or before the hand movement by measuring the mean Euclidian distance of the fixations to the reference point (the object). The higher the mean distance is, the more scattered the fixations would be.

#### *Fixation Dispersion during Hand Movement*

We found no significant effect of *Angle* on the fixations’ dispersion during Hand Movement,  $F(2,28) = 1.262$ ,  $p = .299$ . However we found a significant effect of *Response*,  $F(1,14) = 4.996$ ,  $p = .042$ ,  $\eta^2_p = .26$  indicating that the „correct’ trials had systematic lower dispersion ( $M = 36.04$ ,  $SD = 6.74$ ) than the „incorrect trials ( $M = 38.32$ ,  $SD = 5.79$ ) regardless of the Angle. No interaction effect were found between *Angle* and *Response*,  $F(2,28) = 1.936$ ,  $p = .163$

#### *Fixation Dispersion before Hand Movement*

For the fixations that landed on object before the hand movement, we found an effect of *Angle*,  $F(2,26) = 3.984$ ,  $p = .031$ ,  $\eta^2_p = .23$ . Post-hoc analysis showed that the dispersion is higher on „front’ view ( $M = 51.82$ ,  $SD = 8.55$ ) than on ‘top’ view ( $M = 43.87$ ,  $SD = 10.07$ ),  $p = .041$ , no difference were found with the „lateral’ view ( $M = 49.98$ ,  $SD = 6.55$ ). However, we found no significant effect of *Response*,  $F(1,13) = 1.517$ ,  $p = .240$ . No interaction effect were found between *Angle* and *Response*,  $F(2,26) = 0.452$ ,  $p = .641$  (Figure 5).

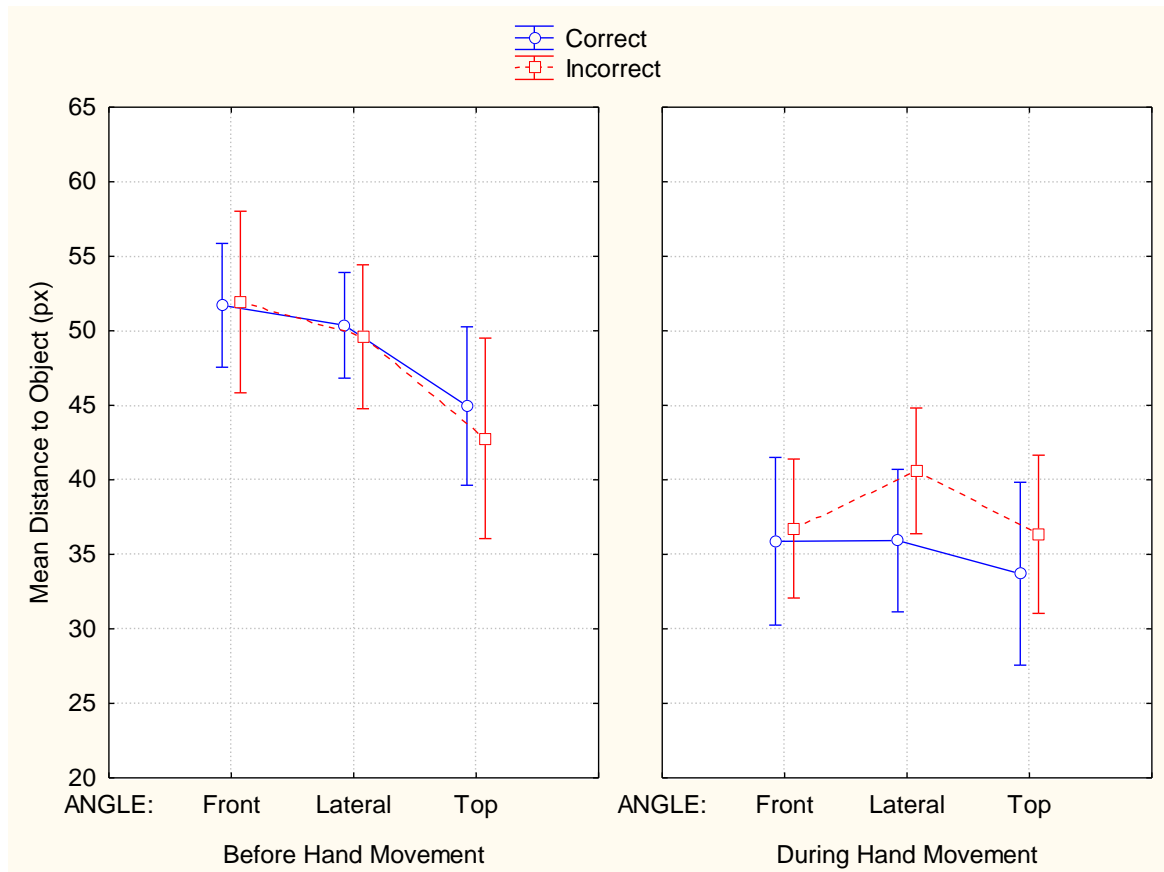


Figure 5. Fixation Dispersion

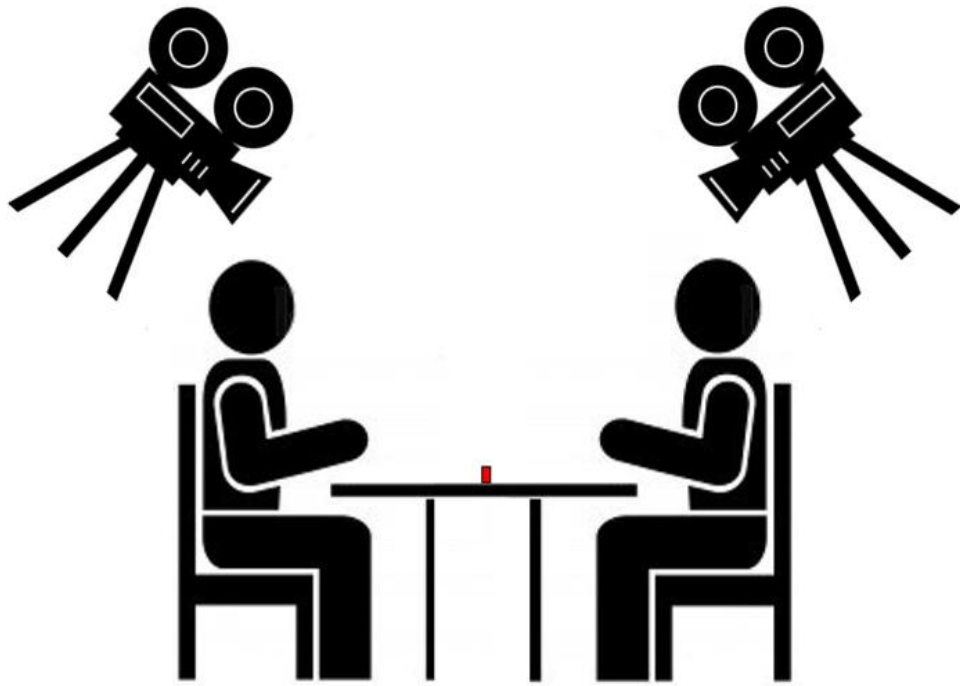
## Conclusion

(NB : These conclusions are partial since more experiments are currently conducted)

Taken together these results suggested that *before* the hand movement, the eye would be guided by the low-level characteristics proprieties of the stimulus. Thus, because the camera’s perspectives modified the basic characteristics of the stimuli, we found a strong *Angle* effect on dispersion *before* the hand movement. More importantly, during the hand movement, no *Angle* effect was reported but a *Response* effect demonstrating that during this part of the

movie, the fixations are less scattered during correct trials than during incorrect trials. More experiments are now needed to ask whether by imposing a specific strategy we could improve the overall classification performance of the participants.





## **Chapitre 5. Interaction Sociale**



# 5.1. Effects of social intention on movement kinematics in cooperative actions

## Abstract:

Optimal control models of biological movements are used to account for those internal variables that constrain voluntary goal-directed actions. They however do not take into account external environmental constraints as those associated to social intention. We investigated here the effects of the social context on kinematic characteristics of sequential actions consisting in placing an object on an initial pad (preparatory action) before reaching and grasping as fast as possible the object to move it to another location (main action). Reach-to-grasp actions were performed either in an isolated condition or in the presence of a partner (*audience effect*), located in the near or far space (*effect of shared reachable space*), and who could intervene on the object in a systematic fashion (*effect of social intention effect*) or not (*effect of social uncertainty*). Results showed an absence of *audience effect* but nevertheless an influence of the social context both on the main and the preparatory actions. In particular, a “localized” *effect of shared reachable space* was observed on the main action, which was smoother when performed within the reachable space of the partner. Furthermore, a “global” effect of *social uncertainty* was observed on both actions with faster and jerkier movements. Finally, *social intention* affected the preparatory action with higher wrist displacements and slower movements when the object was placed for the partner rather than placed for self-use. Overall, these results demonstrate specific effects of action space, social uncertainty and social intention on the planning of reach-to-grasp actions, in particular on the preparatory action, which was performed with no specific execution constraint. These findings underline

the importance of considering the social context in optimal models of action control for human-robot interactions, in particular when focusing on the implementation of motor parameters required to afford intuitive interactions.

## **Introduction**

It's five o'clock and a waiter is faced with the task of clearing a littered table, after a group of customers depart. Through experience, the waiter has learned to produce grip force levels that are adapted to the needs of commonly manipulated objects and to follow hand trajectories that are adapted to the cluttered environment. Empirical studies in laboratory settings have confirmed that physical parameters of an object such as size (Armbrüster & Spijkers, 2006; Chieffi & Gentilucci, 1993; Marteniuk et al., 1990; Pryde, Roy, & Campbell, 1998), weight (Eastough & Edwards, 2007), shape (Gentilucci et al., 1991) and even texture (Fikes, Klatzky, & Lederman, 1994) influence the dynamical aspects of motor performance, in particular the reach-to-grasp motor kinematics. Nevertheless, other internal variables have also been shown to modify motor planning of reaching actions such as the comfort of final posture (David A. Rosenbaum et al., 1990) and the smoothness of movement trajectory (Flash & Hogan, 1985). Most importantly for the matter here, the intention that drives an action can also modulate motor kinematics (Becchio et al., 2008a). Indeed, our waiter may not grasp a glass in the same way if he has the intention to give it to a customer (in this case, the movement may be slow and accurate) or to grip it quickly to put it on a large shelf in order to clean the table before the arrival of the next set of customers. Hence, intention in action as described by Searle (1983) and Jeannerod (2006) represents one category of internal variables that may substantially influence the planning of voluntary action because it encapsulates the fundamental reason of acting.

It is the case that these internal parameters are poorly taken into account in the computational modeling of motor control. Indeed, optimal control models of biological movement are successful in predicting empirical findings such as movement adjustments to

unexpected changes in object position or size, and/or responses to global perturbations (Shadmehr & Mussa-Ivaldi, 1994), and also in modelling the structure of motor variability in function of the physical properties of an object and/or its environment (Gordon, Ghilardi, & Ghez, 1994; Messier & Kalaska, 1999 ; Van Beers, Haggard & Wolpert, 2004) as well as the generic motor laws associated to a given situation (Lacquaniti et al., 1983). However, optimal control models are poorly adapted to predict the empirical data obtained in interactive situations (Friston et al., 2011), rendering human-robot interactions massively unidirectional (Chaminade & Cheng, 2009). Indeed, during social interaction, Boucher et al. (2012) showed that human agents placed in a cooperative context are sensitive to the predictive information provided by the direction of gaze of their partners, even when interacting with robots (Boucher et al., 2012). Furthermore, motor intention influences movement kinematics in such a way that not only the goal of individual actions can be anticipated by a perceiver (Lewkowicz et al., 2013), but also coordinated actions involving several agents can be performed (Knoblich & Sebanz, 2008; Vesper et al., 2010). Thus, it seems important for artificial social intelligence to develop (1) our knowledge of the specific effects that motor intention has on movement kinematics during a true social interactive task and (2) to provide solid guidelines for the development of optimal control models that will be able to implement intention in action in those artificial agents that need to cooperate intuitively with biological organisms.

The effect of motor intention on arm kinematics is a phenomenon that was first reported by Marteniuk and collaborators in the late 1980's (Marteniuk et al., 1987). In this study, they showed that reach-to-grasp movements towards an object differed according to whether the grasped object was afterwards thrown away into a large box or placed into a well. More specifically, results showed that the arm trajectories (i.e., the resultant velocity profile of the wrist) were modulated with an increase in duration of the main deceleration phase of the trajectory when task demands required greater precision. These results did not support a simple scaling procedure in the temporal domain as what would be expected with the optimal

control models of biological movements. Rather, their results supported a view of movement production as relatively specific to the past experiences of the performer and the constraints of the future task. In the continuity of this pioneering study, other studies later reported that not only the final intention but also the characteristics of the second component of a sequential movement could lead to early variants in the first component of the sequence. The effects of a second movement on the first were described in non-manipulative tasks i.e., pointing (J. P. Orliaguet, Vaillon, Coello, & Kandel, 1996) and writing (J.-P. Orliaguet, Kandel, & Boe, 1997). This back propagation effect was also shown in grasping movements when participants were required to grasp (1) an object to eat it or move it (Naish, Reader, Houston-Price, Bremner, & Holmes, 2013), (2) an object to lift or insert it into a niche (Ansuini, Santello, Massaccesi & Castiello, 2006), or (3) a bottle with the intention to use it or to dispose from it (Ansuini et al., 2008; Schuboe, Maldonado, Stork, & Beetz, 2008). More recent studies have finally shown that the final purpose of a grasping action strongly influence the kinematics of both the transport phase and the characteristics of the hand shaping, i.e. the manipulation component (Ansuini et al., 2006). As a consequence, when observing an action performed by someone else, it seems possible from early kinematics to anticipate the goal of the action, i.e. much before the entire action is accomplished (Lewkowicz et al., 2013; Manera et al., 2011; Meary, Chary, Palluel-Germain, & Orliaguet, 2005; Sartori et al., 2011).

Recently, Georgiou and collaborators (Georgiou et al., 2007) showed that the social context while performing a voluntary motor action has also an effect on the kinematics of a reach-to-grasp component of a motor sequence. More specifically, they found that the kinematics of an identical motor action (reaching-to-grasp a wooden block) was different in a cooperative versus a competitive task, and both kinematics patterns could be distinguished from a similar action performed by the participants in isolation. In the same vein, an effect of social intention was reported for movement kinematics when comparing reach-to-grasp actions in a social (passing an object to another person) and a non-social context (putting an object in a concave base, (Becchio et al., 2008a)). Furthermore, social affordances can affect

movement kinematics even when no social interaction is expected (Ferri et al., 2011). In fact, the mere presence of an active conspecific appears sufficient in certain cases to induce changes in movement kinematics (Gianelli, Lugli, Baroni, Nicoletti, & Borghi, 2011). In particular, when participants were requested to grasp an object and then move it to a container, the presence of a person unexpectedly stretching out the arm – as for a social request – affected motor kinematics of those actions that were directed towards the object only (Sartori, Becchio, Bulgheroni, et al., 2009). Interestingly, this pattern of results was not observed when humans interacted with robotic agents, a situation that influenced neither arm trajectories, nor kinematic profiles, suggesting a lack of true social interaction when humans interact with robotic systems. Considered together, these data support the view that specific kinematic patterns characterize and distinguish actions performed in a social and communicative context from those actions executed with a purely individual intent. One reason for this effect of social context on kinematics could be that communicative actions are intended to be identified by a partner and to engage him/her in a communication process (Sartori, Becchio, Bara, et al., 2009). Accordingly, by simply observing the movements performed by others, one might be able to comprehend what they are planning to do and thus, know how one should act in response (Becchio et al., 2012). This point of view fits well with the observation that social effects on reach-to-grasp movement depend on the spatial location of the other person. In particular, latencies in responding have been shown to be significantly shorter when partners are in positions allowing them to easily reach for the object (Gianelli, Scorolli, & Borghi, 2013). Although the presence of another person can influence the latencies and the kinematic profiles of reach to grasp trajectories, specifically when intending to communicate or cooperate with the partner, it is not clear yet whether the social context modulates only those actions that are relevant in the current social situation (reaching, manipulating and displacing objects) or whether the social context modulates all actions that are performed even when they are irrelevant according to the current social and communicative situation.

In the present study, we questioned the specific effect of social intention on movement kinematics for the main manipulative action but also for the preparatory action that was included in the procedure, to initiate each experimental trial. As such, we will be able to discuss whether the social intention induces a general state upon the social behavior or whether social intention has a more specific effect on the action that is carried out towards the target object. To test this hypothesis, participants were asked to reach and grasp as fast as possible an object and to move it to another location. Before performing this main action, participants were required to position the object on an initial pad. In contrast with the main action, this preparatory action was performed without any temporal constraint or direct social interaction. The effects of social context on the kinematic parameters of both the main and the preparatory actions were analysed both when the actions were performed in absence and in the presence of another person, who could intervene on the target object or not depending on his relative position around the workspace and on task instructions.

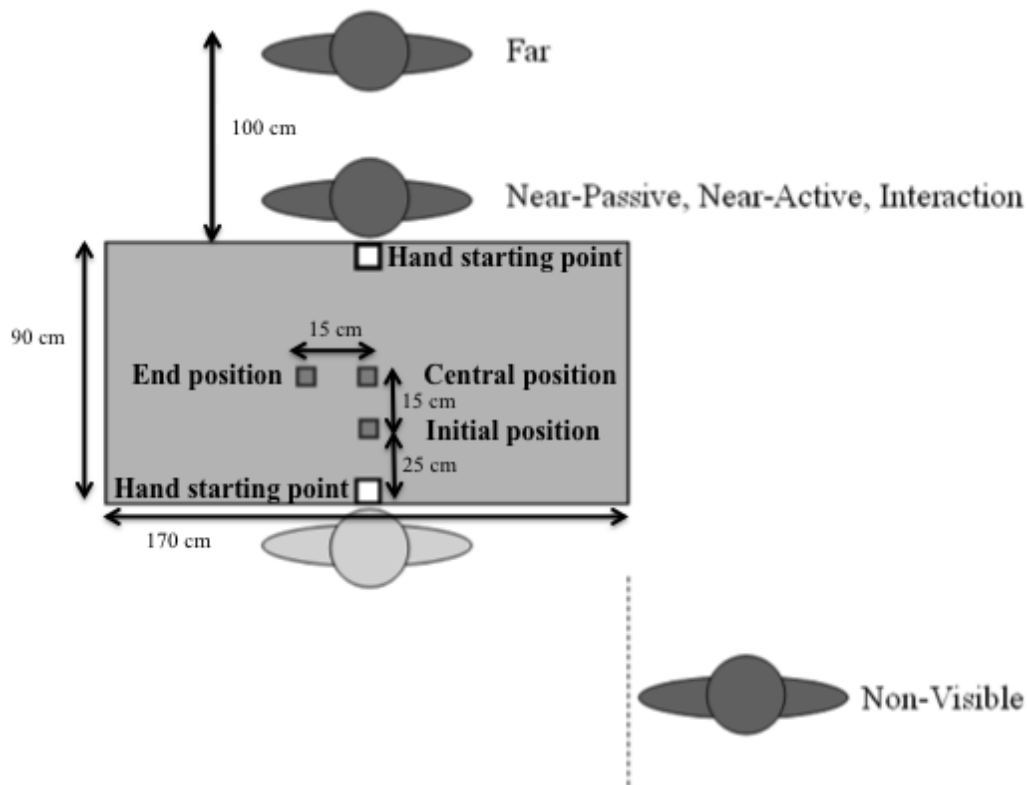
## **Materials and methods**

### *Participants*

Twenty-one healthy adults took part in the experiment (mean age = 22.7, SD = 4.8). All participants were right-handed, with a mean laterality coefficient of 0.88 (Edinburgh Handedness Inventory, Oldfield, 1971) and had no prior knowledge about the scientific aim of the study. Participants provided informed consent before participating in the experiment. The experimenter, a 23-year-old man, played the role of the social partner in all the social conditions requiring a second participant. The protocol followed the general ethics rules defined by the Helsinki guidelines for human experiments and was approved by the local institutional ethic committee.

## Apparatus and stimuli

Participants sat in front of a table (180 x 90 cm) on which red landmarks (3 cm x 3 cm) symbolized three specific locations that will be referred to in the next section as the initial position, the central position and the end position (see Figure 1). In addition, two target-locations were placed on either side of the table, and were used to indicate the starting hand position for both the participant and the experimenter. The object that was to be manipulated was a wooden dowel (width 2 cm and height 4 cm), which was placed on the initial position at the beginning of each trial. In order to prevent any influence of verbal instruction, all trials were triggered through the emission of auditory tones broadcasted by computer speakers.



**Figure 1.** Experimental setup showing the ‘initial’, ‘central’ and ‘end’ positions as well as the respective distances. The position of the participant (light grey) and the partner (dark grey) within the different experimental conditions (absent, near, far) are illustrated. The white squares indicate the starting hand positions for both the participant (*bottom*) and the experimenter (*top*).

## Procedure

During the experimental session, both the participant and the experimenter were seated on either side of the table, facing each other. The participants' task was to reach and grasp the dowel between the thumb and the index finger in order to move it from one position to the other. Each trial started with the object placed at the initial position and with participants pinching their index finger and thumb together, with the fingertips set upon the starting hand position. A trial was defined as a series of three successive action sequences: *Preparatory Action*, which consisted in displacing the dowel from the initial to the central position (no temporal constraints), the *Main Action* which consisted in displacing as fast as possible the dowel from the central to the end position, and the *Repositioning Action* which consisted in displacing the dowel from the central to the initial position (no temporal constraints), making the setup ready for the next trial. Time pressure was set on the *Main Action* only and for this movement, the speed of the participants' wrist was required to be superior to 80% of maximal speed (previously registered, see below). Each move was triggered by a different broadcast tone, which was always played in the same order (tone 1 initiated the *Preparatory Action*; tone 2 initiated the *Main Action*; tone 3 initiated the *Repositioning Action*). In order to prevent participants from anticipating the time of movement initiation, the time intervals between tones were randomized and lasted unpredictably between 1 and 3 seconds.

Tone 2, which initiated the *Main Action*, could be one of two pitches (low or high). When tone 2 was high-pitched, participants were to perform the *Main Action* as quickly and as accurately as possible. When it was a low-pitch tone, participants were to require to refrain from moving and the experimenter was to pick the dowel up from the central position and to place it on the end position as quickly and as accurately as possible.

### *Practice sessions*

All participants underwent two practice blocks before the experimental session started. A first practice block was performed to obtain an estimation of the maximum speed at which



each participant could grasp the wooden dowel from the central position and place it on the end position. We used an adjustment procedure, which consisted in modifying the threshold (maximum speed) according to each participant's performance level. If they were faster than the threshold computed on the last trial, the threshold was increased and reciprocally, if they were slower, it was decreased (by  $50 \text{ mm.s}^{-1}$  at the beginning of the adaptation phase and then, progressively by a smaller change until reaching a  $5 \text{ mm.s}^{-1}$  modulation, at the end of the adaptation phase). The practice block ended when the threshold did not increase or decrease more than three times during the five last consecutive trials, indicating that the threshold was near to the participants' maximum speed. The mean value of the six last measurements was then taken as the individual's speed reference for the *Main Action* in the experimental session. A second practice block (16 trials) was performed in interaction with the experimenter in order to assess whether the instructions were understood by the participants, that the different tones were clearly identified and that the appropriate motor responses were provided.

### *Experimental conditions*

In order to test the contrasting effects of the four different social contexts that were targeted in this study, we designed five experimental conditions in which the experimenter was placed in different places around the table with respect to the participant (see Figure 1). Participants took part in all five conditions following a randomized block design.

Absent. The experimenter was not visible while participants performed the pick and place task. Tone 2 was always a high-pitch sound and thus, all *Main Actions* were performed by the participant.

Far. The experimenter was seated on a chair, facing the participants, at a distance of 100 cm from the table. At the start of the block, the experimenter stretched out his right arm to show the participants that he could not reach the table center. Tone 2 was always a high-pitch sound and thus, all *Main Actions* were performed by the participant.

Near-Passive. The experimenter was seated at the table, facing the participant. At the start of the block, the experimenter stretched out his right limb to show the participants that he could reach the table centre, though he stayed totally immobile throughout the entire experimental session. Tone 2 was always a high-pitch sound and all *Main Actions* were performed by the participant.

Near-Active. The experimenter was seated at the table, facing the participant. At the start of the block, the experimenter stretched out his right arm to show the participants that he could reach the table centre. Tone 2 was always a low-pitch sound and thus, all *Main Actions* were performed by the experimenter.

Interaction. The experimenter was seated at the table, facing the participant. At the start of the block, the experimenter stretched out his right limb to show the participants that he could reach the table centre. Tone 2 was a high-pitch sound in 70% of the *Action* trials and was a low-pitch sound in the remaining 30%. Thus, the *Main Actions* were performed by the participant in 70% of the trials.

A given condition ended when a score of 20 points was achieved. Each point was obtained when a correct *Main Action* was performed, i.e. when the motor performance satisfied the temporal, spatial and social constraints.

### *Data recording and analysis*

The participants' movements were recorded using 4 Oqus infrared cameras (Qualisys system). Kinematics of reach-to-grasp and transport movements were measured by recording the 3D displacement of the 5 infrared reflective markers that were placed on the index (base and tip), the thumb (tip) and the wrist (scaphoid and pisiform) of the participant. One additional marker was placed on the dowel. Cameras were calibrated before each session, allowing the system to reach standard deviation accuracies smaller than 0.2mm, at a 200 Hz sampling rate.

From these measures, tangential 3D instantaneous velocity profiles were calculated. All movements (*Preparatory Action*, *Main Action*, *Repositioning Action*) were characterized by two bell-shaped profiles. The first bell-shape curve corresponded systematically to the movement of reaching to pick the target object, which will be referred to in the following as the first movement of the sequence. The second bell-shape curve corresponded to the movement of lifting to place the target-object, which will be referred to in the following as the second movement of the sequence. For both movements, kinematic parameters of the arm and of the grip components were measured. As classically used in previous studies, reaction time, trajectory amplitude and early kinematic parameters (amplitude and time to peak of acceleration and velocity phases) were here used because they inform on the motor planning properties, whereas movement time and trajectory smoothness (as revealed by jerk analysis) inform on the guiding strategies that are used to displace the hand through action space. These parameters have been pointed out to be relevant indicators for human observers that were required to extract meaningful interaction-cues when viewing point-light displays (e.g., (Cook, Saygin, Swain, & Blakemore, 2009; Pollick, Paterson, Bruderlin, & Sanford, 2001). Definitions and codings of the parameters that we selected in the present study are proposed in Table 1.

NAME	CODE	UNIT	DEFINITION
Reaction time	RT	ms	Time duration between tone onset and first moment in time for which velocity of wrist marker was greater than threshold of 20 mm.s <sup>-1</sup>
Movement time of Mvt <sub>1</sub>	MT <sub>1</sub>	ms	Time duration of the first element of the sequence, equals to the moment in time for which the local minima between the two “bells” occurs minus the reaction time.
Amplitude of peak velocity of Mvt <sub>1</sub>	APV <sub>1</sub>	mm.s <sup>-1</sup>	Amplitude of the first peak of velocity (first zero crossing of acceleration)
Time To Peak Velocity	TPV <sub>1</sub>	ms	Moment in time for which the first peak of velocity occurs minus the reaction time
Amplitude of peak acceleration of Mvt <sub>1</sub>	APA <sub>1</sub>	mm.s <sup>-2</sup>	Amplitude of the maxima of the first derivate of velocity between the start of movement and the peak of velocity.
Mean jerk during acceleration phase of Mvt <sub>1</sub>	Jerk <sub>1</sub>	mm.s <sup>-3</sup>	Mean of absolute values of jerk : second derivate of velocity between the start of the movement and the peak of velocity.
Amplitude of peak height of Mvt <sub>1</sub>	APH <sub>1</sub>	mm	Amplitude of the first maximum value from Z-axis data during element 1.
Maximum grip aperture	MGA	mm	Amplitude of the maximum of the distance between index and thumb marker during element 1.
Time to maximum grip aperture	TGA	ms	Moment in time for which maximum grip aperture occurs.

*MT2, APV2, TPV2, APA2, Jerk2 and APH2 are the same kinematic parameters as above but extracted from Mvt<sub>2</sub> (second bell-shape on velocity profiles for Action trials).*

**Table 1.** Definition of the different kinematic parameters considered in the study.

In the present study, we report the analyses that were conducted on the *Preparatory Action* and *Main Action* only. The *Repositioning Action* was not analyzed. For each participant and condition, the kinematic parameters were submitted to a repeated-measure ANOVA with the 5-level Condition as within factor. The alpha level of significance was set to 0.05. To further investigate the main effect of Condition, we used a posteriori contrasts (see details of matrix coefficients in Table 2). More specifically, we tested the effect of audience by opposing *Absent* against all other conditions ( $\Psi_1$ ). We operationalized the effect of sharing reachable space by opposing *Far* against those conditions for which the experimenter was sitting at the table ( $\Psi_2$ ). We tested the effect of social uncertainty by opposing *Interaction* against the conditions for which there was no ambiguity about who was required to perform the *Main Action* ( $\Psi_3$ ). Finally, for the *Preparatory Action*, we tested the effect of social intention by opposing *Near-passive* and *Near-active* conditions ( $\Psi_4$ ). As these four contrasts are orthogonal, they are independent and will provide the means to assess the explanatory power of each contrast for a given main effect.

---

### **Preparatory Action**

---

Contrast	Non-visible	Far	Near-passive	Near-active	Interaction	$\Sigma Ca$
----------	-------------	-----	--------------	-------------	-------------	-------------

---

$\Psi_1$	+4	-1	-1	-1	-1	0
$\Psi_2$	0	+3	-1	-1	-1	0
$\Psi_3$	0	0	-1	-1	+2	0
$\Psi_4$	0	0	+1	-1	0	0

---



---

### Main Action

Contrast	Non-visible	Far	Near-passive	Near-active	Interaction	$\sum Ca$
$\Psi_1$	+3	-1	-1	0	-1	0
$\Psi_2$	0	+2	-1	0	-1	0
$\Psi_3$	0	0	+1	0	-1	0

**Table 2.** Presentation of the orthogonal post-hoc contrasts that were used to assess the social effects in the *Preparatory Action* (top) and the *Main Action* (bottom), respectively.

## Results

### 1.1 Preparatory Action

Concerning the *Preparatory Action* we observed a global effect of Condition on RT ( $F(4,80)=21.458$ ,  $p<.001$ ,  $\eta^2_p=.52$ ) and TGA ( $F(4,80)=6.548$ ,  $p=.019$ ,  $\eta^2_p=.14$ ). For the first

movement of the sequence, the effects of Condition was also significant on  $MT_1$  ( $F(4,80)=3.257$ ,  $p=.016$ ,  $\eta^2_p=.14$ ),  $TPV_1$  ( $F(4,80)=3.103$ ,  $p=.020$ ,  $\eta^2_p=.13$ ),  $Jerk_1$  ( $F(4,80)=2.579$ ,  $p=.044$ ,  $\eta^2_p=.11$ ),  $APH_1$  ( $F(4,80)=3.317$ ,  $p=.014$ ,  $\eta^2_p=.14$ ). For the second movement of the sequence, the effect of Condition was significant on  $APH_2$  ( $F(4,80)=3.450$ ,  $p=.012$ ,  $\eta^2_p=.15$ ). No effects were found on end-point errors ( $F(4,80)=1.41$ ,  $p=.236$ ), indicating that the end-point accuracy was maintained constant throughout all experimental conditions and thus, did not provide an account for the effects observed on motor kinematics. These results indicate that the presence, the location and/or the interaction with the experimenter were taken into account during motor planning and modulated motor execution. To obtain more specifics about the effects that were impacting movement parameters, we conducted a series of post-hoc contrast analyses.

#### Effect of audience

No kinematic parameters were found to be significantly affected when comparing the Absent condition versus the three other conditions. RT was found to be only close to significance ( $t=1.947$ ,  $p=.065$ ) thus suggesting no audience effect on RT. In agreement with this, we observed an absence of Condition effect on all 16 kinematic parameters, confirming a weak audience effect on motor performances.

#### Effect of sharing reachable space

The results showed an effect of reachable space on RT when contrasting the conditions (Far) and (Near-passive, Near-active and Interaction). Participants performed the *Preparatory Action* with a longer RT ( $t=3.78$ ,  $p=.001$ ) in the Far condition. We also found that the increase of RT in all the Near conditions was the most significant for the Near-active ( $M=410$  ms,  $SD=55$  ms) and Interaction conditions ( $M=360$  ms,  $SD=55$  ms) as compared to the Far condition ( $M=328$  ms,  $SD=60$  ms). No differences were found between Far and Near-passive conditions ( $M=320$  ms,  $SD=46$  ms,  $p=.979$ ), suggesting that the observed effects were supported by other more specific and independent variables (e.g., social interaction). No

effect on MT or kinematic parameters was observed. Thus, we hypothesized that the global effect on kinematics reported above were not due to the near presence of the partner but rather due to the interactive process that take place during the other experimental conditions. To verify this hypothesis, we dissociated two different contrasting hypotheses within the three “Near” conditions. First, we tested the effect of social uncertainty by contrasting (Interaction) versus (Near-passive and Near-active) conditions considered together. Second, we tested the effect of social intention by contrasting the conditions (Near-passive) versus (Near-active).

#### Effect of social uncertainty

When contrasting (Interaction) versus (Near-passive and Near-active) conditions, the results showed an effect of social uncertainty on the kinematic parameters of the first movement of the sequence with shorter  $MT_1$  ( $t=2.756$ ,  $p=.012$ ), shorter  $TPV_1$  ( $t=3.611$ ,  $p=.002$ ), higher  $Jerk_1$  ( $t=2.735$ ,  $p=0.128$ ) and shorter TGA ( $t=2.427$ ,  $p=.025$ ) in the interaction condition compared to the two other conditions considered together. Because all aspects of the task were maintained identical (i.e., starting position, relative positions of participant and experimenter, object location and size, end-position and end-point accuracy) but the social context, the only variable that could account for these results was the uncertainty of whether the next movement would be performed by the participant or the experimenter. Moreover, in the *Preparatory Action* condition the audio stimulus was strictly the same regardless of the condition (near-active, near-passive and interaction). Thus, the effects reported could not be accounted for by a stimulus-response contingency effect but would be more related to the social situation per se.

#### Effect of social intention

When participants initiated the task under the Near-passive condition, results revealed a significant shorter RT ( $t=10.823$ ,  $p<.001$ ) and shorter TGA ( $t=2.727$ ,  $p=.013$ ) than when participants initiated the task under the Near-active condition. For the first movement, a shorter  $MT_1$  ( $t=2.918$ ,  $p=.009$ ), a lower  $APH_1$  ( $t=2.424$ ,  $p=.025$ ) was also observed along with



a lower APH<sub>2</sub> ( $t=2.510$ ,  $p=.021$ ) for the second movement in the Near-passive compared to the Near-active conditions. These results indicate that even though the “motor” intention is the same, the “social” intention involved in the task is taken into account during the planning of the *Preparatory Action*, as reflected in the kinematic parameters of both the first and the second component of the action sequence.

	Preparatory Action						Main Action			
	Absent	Far	Near-passive	Near-active	Interaction		Absent	Far	Near-Active	Interaction
RT	334 (15)	328 (13)	320 (10)	410 (12)	360 (12)	Grasping	210 (6)	220 (8)	226 (6)	267 (9)
TGA	331 (10)	326 (8)	331 (9)	341 (8)	325 (10)		344 (13)	350 (14)	346 (14)	329 (14)
MGA	80.7 (1.7)	80.5 (2.1)	80.9 (2.1)	81.2 (1.9)	79.4 (1.6)		94.4 (3.3)	91.8 (3.1)	92.7 (3.1)	92.9 (3.2)
						Reaching				
APV <sub>1</sub>	710 (22)	717 (19)	711 (23)	708 (19)	721 (27)		1503 (56)	1498 (50)	1504 (54)	1579 (61)
TPV <sub>1</sub>	227 (7)	229 (6)	229 (5)	237 (6)	226 (7)		235 (8)	236 (8)	236 (8)	223 (9)
MT <sub>1</sub>	423 (13)	421 (11)	422 (12)	438 (12)	417 (11)	Placing	447 (12)	450 (14)	445 (14)	431 (15)
APA <sub>1</sub>	5608 (305)	5480 (342)	5662 (321)	5637 (296)	5970 (386)		11171 (853)	10836 (743)	11207 (879)	12511 (1032)
Jerk <sub>1</sub>	3206 (178)	3169 (155)	3159 (158)	3042 (152)	3322 (203)		6507 (453)	6382 (396)	6497 (450)	7235 (517)
APH <sub>1</sub>	61.5 (1.6)	59.8 (1.6)	60.4 (1.7)	62.8 (1.5)	61.1 (1.7)	Placing	68.4 (2.1)	67.1 (2.0)	67.1 (2.1)	68.0 (2.1)
APV <sub>2</sub>	719 (15)	725 (14)	715 (16)	702 (17)	715 (16)		765 (22)	770 (24)	757 (24)	751 (23)
TPV <sub>2</sub>	171 (6)	174 (5)	173 (5)	173 (6)	173 (6)		127 (6)	130 (6)	126 (5)	125 (6)
MT <sub>2</sub>	499 (14)	505 (16)	498 (15)	517 (16)	509 (18)		365 (11)	363 (10)	364 (10)	357 (11)
APA <sub>2</sub>	4132 (229)	4169 (198)	4086 (210)	4059 (184)	4144 (209)		4391 (317)	4507 (282)	4165 (266)	4016 (266)
Jerk <sub>2</sub>	2343 (117)	2373 (106)	2282 (106)	2293 (98)	2357 (114)		2597 (188)	2626 (166)	2442 (163)	2339 (166)
APH <sub>2</sub>	67.7 (2.2)	66.6 (2.2)	67.4 (2.1)	72.6 (2.5)	68.3 (2.6)		74.3 (2.6)	70.9 (1.7)	72.4 (2.1)	73.4 (2.2)

Results are presented for each experimental condition for the Preparatory Action and for the Main Action, respectively.

## 1.2 Main Action

When considering the *Main Action*, the statistical analyses revealed a global effect of Condition on RT ( $F(3,60)=33.806$ ,  $p<.001$ ,  $\eta^2_p=.63$ ) and TGA ( $F(3,60)=6.548$ ,  $p<.001$ ,  $\eta^2_p=.25$ ) as well as on 5 other kinematic parameters characterizing the first movement of the sequence, i.e., APV<sub>1</sub> ( $F(3,60)=7.814$ ,  $p<.001$ ,  $\eta^2_p=.28$ ), TPV<sub>1</sub> ( $F(3,60)=8.690$ ,  $p<.001$ ,  $\eta^2_p=.30$ ), MT<sub>1</sub> ( $F(3,60)=3.827$ ,  $p=.014$ ,  $\eta^2_p=.16$ ), APA<sub>1</sub> ( $F(3,60)=9.076$ ,  $p<.001$ ,  $\eta^2_p=.31$ ), and Jerk<sub>1</sub> ( $F(3,60)=11.397$ ,  $p<.001$ ,  $\eta^2_p=.36$ ). For the second movement of the sequence, results revealed an effect of Condition on APA<sub>2</sub> ( $F(3,60)=3.326$ ,  $p=.026$ ,  $\eta^2_p=.14$ ) and Jerk<sub>2</sub>

( $F(3,60)=3.816$ ,  $p=.014$ ,  $\eta^2_p=.16$ ) only. No effects of Condition were revealed on any of the other kinematic parameters, MGA or end-point errors.

Because all aspects of the task were maintained identical throughout all conditions (i.e., starting position, relative positions of participant and experimenter, object location and size, end-position and end-point accuracy) except for the social context, these findings strongly suggest a global planning of the motor sequences during which the social context is taken into account, with as a consequence the modulation of the kinematic properties of both movements of the action sequence. To gather more information about the specific effects and the role played by the social context on these effects, we conducted a series of post-hoc contrast analyses according to the three hypotheses mentioned above.

#### Effect of audience

When comparing Absent versus the three others conditions, we found an effect of audience on RT ( $t=6.01$ ,  $p<.001$ ). Participants initiated movements faster in the Absent condition ( $M=210$  ms,  $SD=25$  ms) compared to the Far ( $M=220$  ms,  $SD=35$  ms), Near-passive ( $M=226$  ms,  $SD=27$  ms) and Interaction ( $M=267$  ms,  $SD=42$  ms) conditions. The audience effect did not have a significant effect on any other of the kinematic parameters. Overall these findings suggest that, as for the *Preparatory Action*, when taken independently from the other effects (space, uncertainty), the mere presence of a partner had little effect on motor performance.

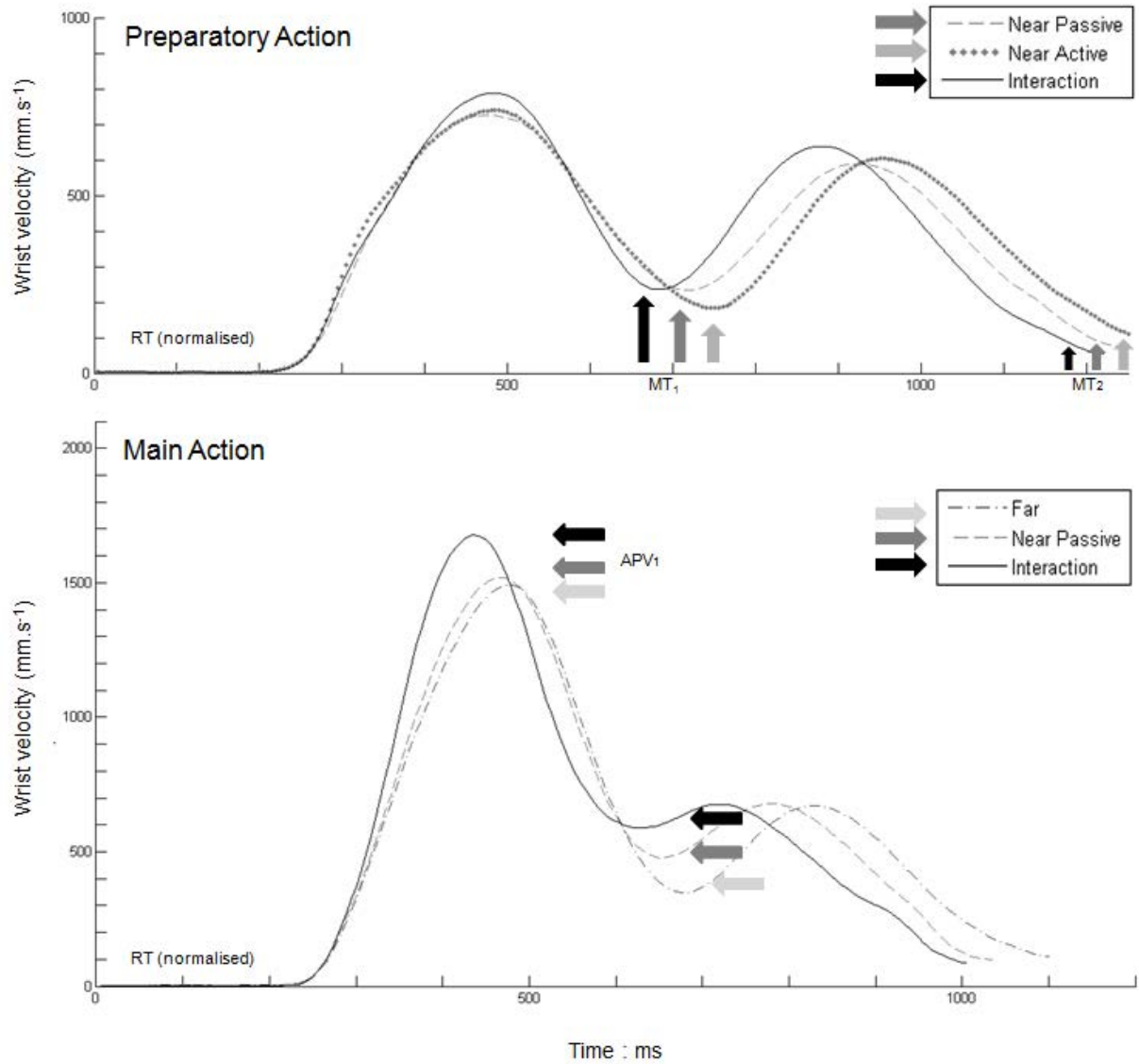
#### Effect of sharing reachable space

When contrasting the conditions (Far) versus (Near-passive and Interaction), statistical analyses revealed that  $APA_2$  ( $t=2.48$ ,  $p=.022$ ) and  $Jerk_2$  ( $t=2.40$ ,  $p=.026$ ) were greater when the partner was far from the participants than when he was near ( $APA_2$ :  $M=4507$  mm.s<sup>-2</sup>,  $SD=1290$  mm.s<sup>-2</sup>;  $Jerk_2$ :  $M=2626$  mm.s<sup>-3</sup>,  $SD=760$  mm.s<sup>-3</sup>). Indeed, both passive ( $APA_2$ :  $M=4165$  mm.s<sup>-2</sup>,  $SD=1220$  mm.s<sup>-2</sup>;  $Jerk_2$ :  $M=2442$  mm.s<sup>-3</sup>,  $SD=748$  mm.s<sup>-3</sup>) and interaction

conditions (APA<sub>2</sub>:  $M=4016 \text{ mm.s}^{-2}$ ,  $SD=1220 \text{ mm.s}^{-2}$ ; Jerk<sub>2</sub>:  $M=2339 \text{ mm.s}^{-3}$ ,  $SD=761 \text{ mm.s}^{-3}$ ) showed small APA<sub>2</sub> and low Jerk<sub>2</sub>, indicating a more fluent transport phase during the sequential action when performed within the partner reachable space.

### Effect of social uncertainty

When contrasting the conditions (Interaction) versus (Near-passive), participants were characterized by longer RT ( $M=267 \text{ ms}$ ,  $SD=42 \text{ ms}$  vs  $M=226 \text{ ms}$ ,  $SD=27 \text{ ms}$ ,  $t=5.44$ ,  $p<.001$ ) and shorter TGA ( $M=329 \text{ ms}$ ,  $SD=65 \text{ ms}$  vs  $M=346 \text{ ms}$ ,  $SD=65 \text{ ms}$ ,  $t=-4.96$ ,  $p<.001$ ). Data analyses also revealed higher APV<sub>1</sub> ( $M=1578 \text{ mm.s}^{-1}$ ,  $SD=280 \text{ mm.s}^{-1}$  vs  $M=1504 \text{ mm.s}^{-1}$ ,  $SD=246 \text{ mm.s}^{-1}$ ,  $t=4.13$ ,  $p<.001$ ), shorter MT<sub>1</sub> ( $M=431 \text{ ms}$ ,  $SD=67 \text{ ms}$  vs  $M=445 \text{ ms}$ ,  $SD=63 \text{ ms}$ ,  $t=3.39$ ,  $p=.003$ ), shorter TPV<sub>1</sub> ( $M=223 \text{ ms}$ ,  $SD=40 \text{ ms}$  vs  $M=236 \text{ ms}$ ,  $SD=38 \text{ ms}$ ,  $t=5.11$ ,  $p<.001$ ), higher APA<sub>1</sub> ( $M=12511 \text{ mm.s}^{-2}$ ,  $SD=4728 \text{ mm.s}^{-2}$  vs  $M=11207 \text{ mm.s}^{-2}$ ,  $SD=4028 \text{ mm.s}^{-2}$ ,  $t=4.53$ ,  $p<.001$ ), and higher Jerk<sub>1</sub> ( $M=7235 \text{ mm.s}^{-3}$ ,  $SD=2367 \text{ mm.s}^{-3}$  vs  $M=6497 \text{ mm.s}^{-3}$ ,  $SD=2064 \text{ mm.s}^{-3}$ ,  $t=5.08$ ,  $p<.001$ ) in the Interaction condition compared to that observed in the Near-passive condition. Furthermore, data analysis testing for the effects of Condition on MT<sub>2</sub> was close to significant ( $t=2.00$ ,  $p=.059$ ) with a tendency for shorter MT<sub>2</sub> ( $M=357 \text{ ms}$ ,  $SD=50 \text{ ms}$  vs  $M=365 \text{ ms}$ ,  $SD=48 \text{ ms}$ ) in the interaction condition compared to that measured in the Near-passive condition. These results suggest a global effect of social uncertainty with longer reaction times and faster and less fluent action execution when acting under the uncertainty that the partner may perform the main action (in 30% of trials). However, these effects were mainly observed on the first movement with little effects on the second.



**Figure 2.** Mean kinematic patterns for a typical participant in the different experimental conditions. All patterns are synchronized to the initiation time. On the *preparatory action* (top), we observed both an effect of social uncertainty (Near Passive and Near Active vs Interaction) and an effect of social intention (Near Passive vs Near Active) on the first movement time. On the *Main Action* (bottom), results showed that accelerative part of the second movement is modified when acting in the partner reachable space (Far vs Near Passive and Interaction). Moreover, social uncertainty (Near Passive vs Interaction) affect strongly the first movement time and the first peak of velocity.

## Discussion

The aim of the present study was to evaluate the influence of reachable space, social uncertainty and social intention on movement kinematics characterizing a sequential

manipulative action that consisted in placing a dowel (preparatory action) before performing a temporally constrained task (main action) that required participants to move as fast as possible the dowel from one location to another. The analyses of the kinematic patterns of both the preparatory (executed under no constraints) and the main action (executed under speeded constraints) revealed an absence of influence of the mere presence of a partner, i.e., the audience effect was negligible. However, there was a significant effect of the social context with variations of movement kinematics of the main action but also of the preparatory action when the partner was located close enough to the table to be able to intervene on the object. Overall, our data suggest a specific effect of the social risk of “sharing reachable space”. In the following sections, we will quickly review the reported results and propose a discussion on the importance of these findings for the field of neuro-robotics.

Using a rather simple reach to grasp task, we manipulated the effect of audience, the effect of sharing reachable space, the effect of social uncertainty and the effect of social intention. First, although it is well established that the mere presence of a partner can affect participants' behaviour (Zajonc, 1965), results showed that the presence of a potential partner was not sufficient to affect the kinematics of the grasping and placing phase of a manipulative task (*Main Action*). These results are in agreement with earlier studies, which reported that movement kinematics is affected by the presence of another person only when an interaction between the two agents can occur (Becchio, Sartori, Bulgheroni, & Castiello, 2008a; Becchio, Sartori, Bulgheroni, & Castiello, 2008b; Georgiou, Becchio, Glover & Castiello, 2007). In contrast, we observed that the presence of a partner sharing the participants' reachable space had a significant effect on the properties of movement kinematics with longer reaction times and lower acceleration peaks, which rendered the arm trajectories less jerky (more fluent). These findings suggest that the presence of a partner sharing reachable space lead the actor to slow down the motor planning process in order to enhance movement guiding strategies, resulting thus in a more fluent transport phase of the sequential action. To note is the fact that these patterns of results were observed essentially for the second element of the main action

(i.e., the transport phase). At first, it may be thought that these results suggest that kinematic modulations were associated to the space variability of the object that is placed on the table. However, through the use of real-time control for small error acceptance, we controlled for this factor: the kinematic variations could not be due to the end-point accuracy constraints and may in fact directly be related to the experimental conditions. In agreement with previous work (Gianelli et al., 2013), these findings indicate that grasping an object to transport it to a new location is affected by whether this object is located in someone else's reachable space, notwithstanding the fact that the aim to interact is made explicit or not. Hence, the fact that movements were smoother and performed with lower acceleration profiles when executed in other's reachable space suggests that grasping actions are influenced by the possibility of experiencing a social interaction.

The main finding of the present study is however the fact that social context influenced not only the kinematics of the main action but also the kinematics of the preparatory action for which no instructions were given for temporal, spatial or social constraints and despite the fact that this movement was entirely performed out of the reachable space of the partner. Overall, we report in the present study similar effects of social uncertainty in both the *Preparatory Action* and the *Main Action*, showing that the interaction condition not only influenced the grasping task performed as fast as possible (in order to be rewarded by points), but also the preparatory sequence of this action, which was performed always by the participant. Social uncertainty led participants to perform faster preparatory actions, resulting in earlier time to peak velocity and grasp aperture as well as increased jerk. These results indicate that participants felt an urge to perform the preparatory action with shorter response times when the experimental condition generated ambiguity about who will then act. Indeed, in the interaction condition, during the preparatory actions, participants did not know who was going to perform the main action since the sound indicating the agent was given after the preparatory action had been executed. Hence, social uncertainty led participants to adopt a general competitive behavior, which has been previously described in

paradigms that are however usually designed specifically to encourage direct competition (Georgiou, Becchio, Glover & Castiello, 2007). Effect of the social context on movement kinematics was also observed during actions for which the object was placed in totally predictive contexts: data showed that participants tended to have longer reaction times and movement times, and performed more curved trajectories (e.g., higher wrist displacement, APH) when they positioned the object for a forthcoming social action performed by a partner in 100% of the trials (Near passive condition) rather than individually (Near active condition). Variations in movement kinematics are then observed when participants place the object knowing that the partner is going to grasp it and when they place the object knowing that they will personally have to grasp it. Slower actions and higher wrist trajectories may have been implemented to attract the partner's attention and give the person time to prepare the interactive response (Sartori, Becchio, Bara & Castiello, 2009). Interestingly, this pattern of results was obtained even if the motor intention was identical throughout, i.e., an identical target and a similar motor task. It confirms the influence of social intention on movement kinematics, as already reported by Becchio, Sartori, Bulgheroni and Castiello (2008b), and further demonstrates the effect of social intention on motor behaviors as a global effect that affects both the early and the late execution of a motor sequence, including both a preparatory and main action.

Our results reinforce the importance in computational models of motor control to take into account the contextual constraints such as reachable space, environment predictability and social intentions. Current models of motor control (for a review, see Todorov, 2004) are based on optimized function costs that are often named minimum-X (jerk, torque change, energy, time, variance, etc). The present results further demonstrate that such optimized function cannot account for the specific effects that we have reported both for the main and the preparatory actions. Here, we confirm that interacting with a partner encompass different processes that may be independent from each other. First, the effect of reachable space was found to be a “localized” effect on kinematics only revealed when the movement was directly

made within the reachable space of the conspecific. The observed consequence is that the accelerations ( $APA_2$  -  $jerk_2$ ) were reduced giving rise to smoother movements. This could be a consequence of years of learning that when acting within the reachable space of someone else, the agent must have smoother movements in order to not frighten the partner away, smoother and slower profiles being perceived as more gentle and socially engaging actions. This specific learning could be shaped during the early developmental years when young children are interacting with their parents, individuals who are there to teach how to “be gentle” during social interactions (Gaussier et al., 1998; Hasnain, Gaussier, & Mostafaoui, 2012). Second, the effect of social uncertainty is found to be a “global” effect on kinematics, neither localized to a specific part of the sequence, nor at a specific spatial location between the participant and the partner. We found that when the agent cannot entirely predict who will perform the next Main Action, (s)he was performing voluntary actions as if they were in a competitive interaction and thus, modulated both the first and the second components of the motor sequence resulting in less smooth movements (higher accelerations; higher jerk). More experiments are now needed to better understand how the perception of a competition situation in relation to the social context may influence the kinematics of voluntary motor actions. Third, we revealed an effect of social intentions independently from the previous effects. In our case, the preparatory action showed specific patterns of movement curvature with higher wrist displacements and slower movements when participants placed the object to be grasped by the partner compared to the situation for which the object was placed for self-use. Because this situation led to a less “optimized” motor performance, one may speculate that this strategy would be employed as an external signal during social interaction to show the agent’s social intention to share the object (Sartori, Becchio, Bara & Castiello, 2009). Previous studies have supported this interpretation by showing that humans are sensitive to external kinematic characteristics of a movement, and especially trajectory height (Sartori et al., 2011; Manera et al., 2011; Lewkowicz et al., 2013). The new findings reported here suggest that even the preparatory actions reflect the agent’s social intention and thus,



movement properties may be *read* by perceivers for whom understanding motor intention from early kinematics is important. This is at least one of the key elements lacking today in humanoid robot systems because they are not implemented at the moment with the appropriate embedded perceptual system that can take advantage of these early motor information.

In conclusion, the present study provides the first report of a social effect on kinematics of a non-constraint action. To summarize, we found that the mere presence of a conspecific did not influence the preparatory action, even when sharing reachable space with that of the actor, but an overall effect was observed when the task involved social uncertainty and social intention. This result is important as it shows that social uncertainty and intentionality influence kinematics very early on during motor planning, and may thus represent a highly informative signal in the case of cooperative and competitive social situations (see also Manera et al., 2011). These empirical results can have significant impact in the field of neuro-robotics as they suggest that acting in a social interactive environment leads to a certain number of parameters that impact movement kinematics directly: reachable space, uncertainty and social intention. These effects may constitute what humans perceive as a “social interactive” situation, effects that need to be taken into account to create robots with what is called today as intuitive interactivity. Future studies need now to consider how to implement these social aspects of motor control within an artificial system in order to afford intention reading during human-robot collaborative work. More specifically, the questions of low/high-level kinematics and explicit/implicit learning will be the key to implement intuitive social interactive capacities in future humanoid robots.



# Chapitre 6. Discussion Générale

## 6.1. Rappel des principaux résultats

Au cours des expériences que nous avons réalisées, nous avons tout d'abord pu montrer qu'à l'aide d'une situation interactive en temps réelle, il était possible d'observer des modifications systématiques dans la planification des actions séquentielles, préalablement à leur exécution. Par exemple, dans le cas de fortes contraintes (vitesse rapide et précision élevée) le phénomène de « planification couplée » fut observé. A la différence d'un modèle de type étape-par-étape, ce type de planification intègre les éléments moteurs de la séquence dans une seule représentation prise comme un tout global. Ainsi, les contraintes qui s'appliquent sur un des éléments de la séquence (la position terminale) vont se propager sur les autres éléments générant des modifications observables dans la cinématique précoce des mouvements. Ces modifications sont suffisamment précoces pour permettre une lecture de l'objectif moteur distal (l'endroit où l'objet sera posé) en observant la manière dont est réalisé l'objectif proximal (comment l'objet est saisi). Dans le cas particulier de notre expérience (Chapitre 3), les positions initiales et terminales étant conservées, lors de l'exécution, nous proposons que la planification de la séquence peut se faire à travers une représentation motrice (ou modèle global) qui tient compte non seulement de la position d'arrivée souhaitée mais aussi des contraintes spatiales et temporelles selon lesquelles la séquence est exécutée. Une extension de ces résultats fut obtenue dans la deuxième série d'expérience pour des vitesses non contraintes et des paramètres contextuels plus difficiles à appréhender que la simple précision terminale. Ainsi, en modifiant la stabilité et la prédictibilité du point terminal de la séquence, nous avons pu mesurer l'apparition de corrélations positives et significatives lorsque les sujets effectuaient une planification globale de la tâche (Chapitre 3.2). D'autres expériences sont en cours pour explorer par exemple comment encore étendre ces résultats pour mieux comprendre la distinction entre la stabilité terminale d'une part et l'incertitude de l'environnement d'autre part (résultats non présentés dans ce manuscrit). Nos résultats

préliminaires semblent indiquer que dans le cas d'une *instabilité* terminale provoquée par exemple par une surface d'arrivée compliant ou élastique, les modifications cinématiques ne se propagent pas sur toute la séquence mais vont modifier uniquement le deuxième élément moteur (déplacement et pose de l'objet). Nous proposons dans ce cas qu'il s'agit de modulations locales pour permettre d'assurer une pose correcte de l'objet. En revanche, en modifiant la *prédictibilité* de l'évènement rotatoire, comme par exemple le fait que la rotation de la plateforme d'arrivée survienne à intervalle régulier (toutes les secondes) ou de manière aléatoire (entre 0.5 et 1.5s), cela nous a permis d'observer des modulations globales sur la séquence qui affectent à la fois la pose de l'objet mais aussi sa prise.

Nos résultats nous permettent de suggérer qu'en fonction des situations environnementales, différents modes de planification vont engendrer des modulations locales ou globales selon la manière dont les contraintes de la tâche sont incorporées lors de la sélection des commandes et des étapes de planification motrice. La nouveauté de nos résultats est la possibilité 1) d'étendre la loi de Fitts (relation entre vitesse et précision) dans une tâche séquentielle de *manipulation d'objet*, qui n'avait été retrouvée jusqu'à présent que pour des mouvements séquentiels de *pointage* et 2) une meilleure définition des conditions nécessaires pour observer un couplage entre les deux éléments moteurs au sein d'une même séquence. Plus précisément, sur ce dernier point il s'agit de nécessité de la continuité spatiale et temporelle entre les deux éléments moteurs suggérant qu'il est possible d'incorporer les actions uniquement si elles sont prédictibles dans le temps et l'espace. Mais par ailleurs, que l'intensité des contraintes (vitesse, précision, etc...) serait à l'origine de la différenciation observée entre les différents modes de planification.

Ainsi, dans des tâches purement individuelles ne mettant aucunement en jeu le problème de l'interaction sociale, nous avons pu démontrer l'existence de modulations cinématiques importantes dans les mouvements d'atteinte et de saisie et donc utilisables précocement du point de vue de l'observateur.

Ces résultats sont en accord avec les résultats obtenus en interaction sociale (Becchio et al. 2008, 2012) mais viennent donner des éléments de réflexion nouveaux quant à l'interprétation des modulations provoquées par l'intention de *donner* l'objet au partenaire. En effet, nous proposons que l'hypothèse classique d'une plus grande précision terminale proposée notamment par l'équipe de Becchio et collaborateurs est insuffisante pour expliquer pourquoi les sujets ont par exemple des pics de vitesse diminués et une plus longue phase de décélération lorsqu'ils donnent l'objet dans la main du partenaire par rapport à lorsqu'ils le posent sur un réceptacle solide (Becchio et al. 2008). Nos résultats semblent suggérer que des modulations semblables s'observent aussi dans un cas purement individuel si la surface du réceptacle est élastique par rapport à un réceptacle solide ou compliant. De plus, nous proposons que la présence de la main du partenaire comme réceptacle engendre en revanche une *prédictibilité* amoindrie et qu'elle serait à l'origine des modulations observées précocement dans la phase d'atteinte et de saisie de l'objet. Cette prédictibilité amoindrie serait due au fait que la main n'est pas considérée comme un réceptacle immobile et qu'il est toujours possible que le partenaire effectue des mouvements non prévus au cours de l'exécution de la séquence. Cependant, cet effet de prédictibilité doit être pris de manière différente par rapport aux effets des contraintes de précision car bien qu'il engendre des modulations dites « globales » qui apparaissent à la fois sur le premier et le deuxième élément moteur, cet effet provoque aussi des dé-corrélations dans la séquence (Chapitre 3.2) par rapport à une situation où tout est prédictible (Chapitre 3.1, expérience 1 et 2).

Par ailleurs, les résultats de l'analyse cinématique des trois catégories de séquences exécutées à vitesse préférée montrent qu'il y ait une forte redondance parmi les caractéristiques cinématiques des mouvements d'atteinte et de saisie (Chapitre 4.1 Partie A). Nous suggérons alors que l'origine de cette redondance serait bien un paramètre d'optimisation au niveau proximal afin d'assurer une bonne prise de l'objet (M. Jeannerod et al., 1995; M. Jeannerod, 1984). Par exemple, la distance à parcourir et les dimensions de l'objet pourraient être des paramètres privilégiés pour sélectionner les commandes motrices

du geste à réaliser (Bootsma et al., 1994; Gentilucci et al., 1997). Cependant, chez l'homme d'autres paramètres, contextuels ou distaux semblent aussi combinés à ce modèle proximal, générant ainsi des déviations par rapport au modèle « optimal » d'un geste d'atteinte seul. Ces résultats ne sont toutefois pas en désaccord complet avec des modèles optimaux qui prennent en compte plusieurs objectifs tels que le minimum-jerk contraint (Todorov et Jordan 2002) qui semble bien adapté pour réaliser des transitions « biologiques » des actions séquentielles. Une des questions à résoudre reste dans quelle mesure plusieurs buts peuvent être incorporés dans une seule séquence, et pourquoi certains sous-buts sont anticipés alors que d'autres sont pris séparément. Nos résultats obtenus dans le chapitre 3.1 semblent indiquer que l'intensité de contraintes qui s'applique sur la séquence serait à l'origine du changement du modèle de planification. Ces résultats sont en faveur de l'hypothèse selon laquelle, le contrôleur optimiserait aussi la boucle fermée (Guigon, Baraduc, & Desmurget, 2003, 2008) c'est-à-dire la manière dont les retours sensoriels sont utilisés par le contrôleur pour la sélection de nouvelles commandes. Ceci permettrait de générer des planifications différentes mais sous-entend une certaine connaissance (explicite ou non) du contexte de la tâche et des contraintes qui s'y appliquent, ce qui n'est pas toujours le cas en situation naturelle. Dans notre situation expérimentale le fait d'agir à la vitesse la plus rapide possible constitue en effet une connaissance préalable d'une des contraintes très fortes de la tâche, qui agit comme un sélecteur de plus haut niveau concernant les choix réalisés par le contrôleur. Ainsi, le participant pourrait sélectionner un mode de contrôle, qui « n'attend pas » les retours sensoriels car les délais seraient trop long pour effectuer des corrections en ligne. En conséquence les mouvements sont beaucoup plus rapides et de fortes erreurs terminales sont observées. Les modèles optimaux s'avèrent ainsi très efficaces pour prédire les cinématiques biologiques observées chez l'humain. Ils parviennent à trouver des solutions pour les problèmes de redondance (degrés de liberté) et peuvent s'apparenter à des inférences de type bayésiennes (voir : Todorov 2009). Mais une des limites de ce type de modèle est toutefois qu'il ne font pas de distinction entre les problèmes de traitements sensoriels, de contrôle

moteur ou de planification motrice (Toussaint, 2009). Ces modèles ne peuvent par exemple pas expliquer pourquoi certains paramètres auront des effets locaux alors que d'autres auront des effets globaux en particulier si l'objectif du modélisateur est de chercher un paramètre unique d'optimisation. Nos résultats peuvent donc servir de suggestions pour développer de nouveaux modèles de contrôle du mouvement qui incorporent les effets des contraintes à la fois proximales et distales dans le but de réaliser des séquences de mouvements biologiques observables qui pourraient alors être anticipées par un humain.

Dans une nouvelle série d'études consacrées à la l'observation et la compréhension des actions. Nous avons pu montrer que les participants adultes et sains étaient capables de catégoriser trois intentions motrices différentes à partir de la simple observation de l'élément moteur précoce, le mouvement d'atteinte et de saisie de l'objet. La première constatation que nous avons fait a été d'observer que les participants n'étaient pas capables de s'auto-évaluer à la hauteur de leurs performances réelles. Autrement dit, bien que cela ne constitue pas une preuve que ce processus est purement automatique, ces résultats suggèrent que nos capacités d'anticipation motrices semblent meilleures que celles qui nous apparaissent spontanément. Par ailleurs, une première modélisation à l'aide d'un système de classification par réseaux de neurones artificiels fut testée pour comparer avec les performances humaines. Au travers de la conception de ce modèle nous avons pu poser une question supplémentaire sur la quantité d'information nécessaire pour discriminer des actions séquentielles avec une architecture qui ne possède pas de fonction cognitive complexe. Nous avons alors montré que les pourcentages corrects de discrimination commençaient à dépasser le niveau de la chance peu avant la fin du geste d'atteinte et de saisie (Chapitre 4.1 Partie C). Avant cette étape cruciale, où les doigts se posent sur l'objet, le classifieur n'est pas capable de prédire l'objectif moteur de la séquence. Ces résultats sont bien en désaccord avec l'hypothèse de Jacob et Jeannerod selon laquelle il ne serait pas possible de prédire l'intention sociale (où l'objet sera placé) si l'intention motrice est maintenue constante (où l'objet est saisi). Cependant, une des limites de ces résultats est le fait que les cinématiques obtenues pour le geste d'atteinte et de saisie

sont belles et bien différentes en fonction des catégories et que cela pourrait être dû à une position finale de la séquence qui n'est pas identique dans les trois catégories.

Toutefois, ces résultats constituent un premier pas vers la possibilité d'une détection précoce (bien avant la fin de la séquence) et sont intéressants pour une architecture artificielle dont l'objectif est de pouvoir anticiper le but final d'une séquence afin d'assister un agent dans ses mouvements. Cependant, il est aussi à noter que le fait de pouvoir catégoriser des actions dans un choix forcé à 3 possibilités est différent de pouvoir véritablement « prédire l'intention » de l'agent. En effet, le réductionnisme expérimental dont nous nous servons pour contrôler correctement nos variables en psychologie ne nous permet pas toujours de pouvoir vérifier ce qu'il se passerait si on ne donnait aucune *proposition* de réponse à nos participants. Pourtant, dans notre vie quotidienne, nous faisons librement des inférences sur les intentions des autres qui de plus peuvent s'avérer fausses en réalité. La question qui reste encore à explorer est, est-ce que nos inférences sont fausses parce que nous n'avons pas correctement exploité un indice externe (ex : cinématique) qui aurait dû être visible ou est-ce que l'indice n'étant pas présent ni exploitable, nous nous sommes basés sur d'autres propriétés comme le contexte de l'action et nos connaissances préalables de la situation ou de la personne observée mais que cela fut insuffisant pour nous donner une idée précise de l'intention de l'autre ? Il serait intéressant de vérifier dans quelle mesure les sujets sont capables d'utiliser leurs capacités d'anticipation dans des situations naturelles comme par exemple en utilisant des petits films à la première personne comme c'est par exemple le cas dans les expériences utilisant des dispositifs d'eye-tracking portables (voir (Hayhoe & Ballard, 2005).

Nous résultats nous permettent de suggérer que les modulations cinématiques qui, mêmes fines et difficilement perceptibles consciemment puissent être à l'origine de la capacité de lecture des intentions chez l'humain. Ces résultats sont en accord avec la littérature concernant le rôle de la cinématique pour prédire les actions (Aglioti, Cesari, Romani, & Urgesi, 2008; Graf et al., 2007; Kandel, Orliaguet, & Boe, 2000; Manera et al., 2010; Martel, Bidet-Ildei, & Coello, 2011; Meary et al., 2005; J. P. Orliaguet et al., 1996;



Pozzo et al., 2006; Sartori et al., 2011; Saunier, Papaxanthis, Vargas, & Pozzo, 2008; Stadler, Springer, Parkinson, & Prinz, 2012; Stapel et al., 2012). Toutefois, étant donné le relatif faible score (très loin du 100% correct), il semblerait que des systèmes *complémentaires* basés sur d'autres indices puissent aider dans la compréhension des actions comme la position du regard du partenaire, les contraintes de la tâche et le contexte environnemental dans lequel ces actions sont entreprises qui déterminent à la fois leur exécution mais aussi l'étendue des actions éventuellement possibles (Brass, Schmitt, Spengler, & Gergely, 2007; Csibra, 2008).

Dans une quatrième série d'études, nous nous sommes intéressés aux mouvements oculaires lors de l'observation d'éléments moteurs précoces (atteinte et saisie) et nous avons pu vérifier que les patterns de mouvements des yeux étaient comparables à ceux réalisés lors de l'exécution de telles séquences. Par ailleurs, nous avons aussi pu vérifier que le taux de mauvaises réponses en catégorisation des intentions pouvait être dû soit à un problème de *qualité* de source d'information soit à un problème de *stratégie* de prise d'information. L'angle de vue de la caméra va ainsi modifier la qualité de l'information du mouvement ce qui se retrouve sur les pourcentages de bonnes réponses. De plus nous avons observé des saccades plus précoces et des fixation plus longues sur l'objet dans le cas de la vue à la première personne (de dessus) en comparaison avec des vues latérales ou de face. Nous suggérons que dans ce cas ce sont les propriétés de bas niveau des vidéos qui vont guider les saccades et les durées de fixation préalablement au déplacement de la main. Cette modification de la *qualité* de l'information peut être renforcée par la possibilité d'une métrique des saccades en amplitude et en vitesse différente selon les points de vue (résultats en cours d'analyse) ce qui pourrait engendrer des biais dans l'estimation des distances si cette estimation était basée sur la décharge corolaire aux mouvement des yeux, par exemple.

Par ailleurs, en comparant les mouvements des yeux lors des réponses correctes par rapport aux réponses incorrectes nous avons trouvé que le *nombre de saccades* préalablement au déplacement de la main était réduit, ce qui pourrait indiquer une meilleure stabilité générale du regard. De plus, nous avons trouvé que la *dispersion des fixations* sur l'objet

pendant le déplacement de la main était plus élevée dans le cas des mauvaises anticipations. Ce dernier résultat indique que le point d'arrivée de la saccade est plus variable lorsque les sujets donnent des mauvaises réponses, ce qui suggère bien que la stratégie de prise d'information peut influencer le pourcentage de bonne réponses. Autrement dit, il semblerait que pour pouvoir discerner et anticiper correctement l'action qui va suivre, les mouvements oculaires proactifs (Johansson et al., 2001; Neggers & Bekkering, 2000) doivent être plus précis et plus stables. Ces résultats sont aussi en accord avec l'idée que l'observation des mouvements de la main entraîne des mouvements oculaires similaires à ceux effectués pendant l'exécution (Flanagan & Johansson, 2003; Rotman et al., 2006). L'intérêt de cette étude est qu'elle vient étendre ces résultats dans une situation d'anticipation motrice où les taux de bonnes réponses semblent corrélés avec une stratégie visuelle proactive plus stable et plus précise et soulignant aussi le fait que l'œil est maintenu sur la cible jusqu'à ce que la main l'atteigne (Neggers & Bekkering, 2000). De plus il semblerait qu'il existe des liens entre importants entre les intentions et la stratégie d'exploration visuelle (Bekkering & Neggers, 2002) avec notamment l'implication du cortex pariétal connu pour ses fonctions de transformation de coordonnées (Buneo & Andersen, 2006; Crawford, Medendorp, & Marotta, 2004). Une extension de ces résultats préliminaires est déjà envisagée concernant le fait d'imposer ou de biaiser le comportement oculaire des participants afin de vérifier l'influence par exemple de la variabilité terminale des saccades sur les capacités d'anticipation motrices.

Enfin, dans une cinquième série d'étude nous avons cherché à vérifier quels effets pourraient être impliqués dans la distinction entre une situation individuelle et une situation sociale (Chapitre 5). Cependant, il est à noter que parmi les paramètres utilisés dans l'expérience (audience, espace d'action, incertitude et intention), au moins deux peuvent aussi être retrouvées dans des situations individuelles (espace et incertitude) et donc ne serait pas forcément suffisants pour décrire une interaction sociale. Par exemple, le problème de l'incertitude peut être relié à un problème de prédictibilité général de l'environnement qui dans le cas d'un environnement non prédictible engendre des vitesses d'exécution plus

rapides au niveau global sur les deux éléments de la séquence. D'autre part, nos résultats montrent que l'influence de l'espace d'action d'autrui engendre uniquement des modifications très locales. Pour certains auteurs, ces résultats confirment l'idée qu'une action potentielle du partenaire va moduler les affordances des objets situés dans l'espace péri-personnel (Costantini, Ambrosini, Scorolli, & Borghi, 2011; Gianelli et al., 2013). Costantini et collaborateurs, ont mesuré des effets de compatibilité stimulus-réponse sur les temps de réactions et proposent l'hypothèse d'une modulation de l'affordance de l'objet selon sa présence dans ou hors de l'espace peri-personnel (Costantini, Ambrosini, et al., 2011). Ils ont ensuite étendus ces résultats à une situation où un avatar en trois-dimensions était proche de l'objet à saisir. De façon intéressante, ils ont retrouvé les effets de compatibilité sur l'objet alors qu'il était loin pour le participant mais dans l'espace du partenaire. De plus, si le mouvement du partenaire est impossible à cause de la présence d'une vitre entre l'avatar et l'objet, les participants ne présentaient plus d'effet de compatibilité stimulus-réponse, suggérant la possibilité d'étendre la notion d'espace péri-personnel à celui d'un partenaire (Costantini, Committeri, & Sinigaglia, 2011). Nos résultats ont montré des modifications cinématique dans la séquence exécutée à vitesse rapide, et uniquement dans la deuxième partie du mouvement, qui s'effectue bien dans l'espace d'action du partenaire. Ces modulations sont dans le sens d'une meilleure fluidité des mouvements si le partenaire est présent avec une réduction du pic d'accélération et du jerk moyen. L'hypothèse de Costantini proposant une modification de la représentation au niveau perceptuel semble difficilement expliquer nos résultats étant donné que le geste d'atteinte en direction de l'objet n'est pas modifié. En revanche, ces modulations observées au niveau cinématique sont similaires à celles observées dans le fait de poser l'objet sur une texture élastique au niveau terminal comparativement à une texture compliant (c'est-à-dire un effet localisé sur la deuxième partie de la séquence). Autrement dit, bien qu'il s'agisse d'un véritable effet de la présence proche du partenaire, les modulations pourraient s'expliquer par des processus de traitement sensori-moteurs comparables à ceux utilisés dans des actions individuelles. De nouvelles

expériences sont nécessaires pour explorer s'il s'agit effectivement des mêmes type de modulation ou non.

Toutefois, le fait que certains paramètres puissent être retrouvés dans des situations individuelles suggère qu'il est possible que des apprentissages moteurs individuels puisse aider le traitement sensori-moteur d'une situation sociale. Nous suggérons ainsi qu'il existe un continuum entre les situations individuelles et les situations sociales. Nous ne pensons pas qu'il existe plusieurs systèmes perceptivo-moteurs (Pierno & Ansuini, 2007), un qui soit spécifique des situations individuelles et un autre spécifique des situations sociales. Bien que des processus complémentaires puissent être mis en jeu, et que les réseaux de neurones dans certaines aires cérébrales puissent avoir développé des spécificités (discuté dans les paragraphes suivants) nous proposons qu'un cadre théorique unifié puisse englober les résultats que nous avons obtenu à la fois avec ou sans présence du partenaire.

En dernier lieu et pour préciser la portée de nos résultats, nous suggérons que *l'intention d'interagir* avec le partenaire pourrait être un des éléments constituant d'une interaction sociale qui n'est pas présent dans une situation individuelle. Ainsi, la manière d'exécuter des mouvements avec l'intention *d'affecter le comportement d'autrui* pourrait avoir été apprise au travers de nombreuses interactions sociales. Nos résultats montrent que l'agent va présenter des variations cinématiques caractéristiques par rapport à un modèle « optimal » de manière à élever le point maximal de hauteur de la trajectoire au cours de la séquence. Nous suggérons alors l'hypothèse selon laquelle cette déviation par rapport à un modèle optimal pourrait être utilisé comme un signal de communication générant une *erreur de prédiction* chez l'observateur. Ce signal pourrait provoquer un effet de « surprise » chez l'observateur afin qu'il soit par exemple plus attentif et qu'il puisse percevoir correctement les mouvements dont l'agent a *l'intention* qu'ils soient lus et interprétés. De plus cela pourrait faciliter l'observation en mettant le point maximal de hauteur (corrélé avec le pic de vitesse) plus haut et plus proche de la hauteur des yeux du partenaire. Cela aura aussi comme conséquence d'augmenter la courbure de la trajectoire ce qui a pour conséquence de « lisser »

les profils de vitesse afin de rendre les mouvements plus fluides et plus prédictibles. Nos résultats semblent donc en accord avec l'hypothèse selon laquelle dans le cas d'actions

Pour résumer, la nouveauté de nos résultats expérimentaux est que le phénomène classiquement appelé « interaction sociale » dans le cadre du contrôle moteur ne semble pas se résumer à l'optimisation d'un paramètre unique tel que la précision terminale ou la variabilité globale du mouvement. Nos résultats montrent qu'une « interaction sociale » serait constitué d'au moins trois entités qui possèdent chacune des influences indépendantes sur la programmation et la planification motrice. Ainsi, la *présence* ou non de l'espace d'action du partenaire, *l'incertitude* lié à la prise d'initiative et enfin *l'intention* d'affecter le comportement du partenaire sont autant de paramètres qui modifient la cinématique des mouvements de l'agent. De nouvelles expériences sont nécessaires pour mesurer précisément les contributions relatives de ces paramètres mais nos résultats constituent un premier pas vers une description plus formelle de l'interaction sociale pour les futurs modèles computationnels des actions séquentielles chez

conjointes, les mouvements réalisés sont plus prédictibles (Vesper, van der Wel, Knoblich, & Sebanz, 2011). Cependant, nos résultats sont limités par le fait que nous avons contrôlé expérimentalement la précision terminale. Il serait intéressant de réaliser des expériences complémentaires en retirant la contrainte de précision pour le mouvement préparatoire par exemple pour vérifier l'effet de l'intention d'affecter le comportement d'autrui sur la variabilité à la fois durant les trajectoires et au niveau des positions terminales.

## 6.2. Conséquences pour les interactions sociales

Dans le cas des actions conjointes (Sebanz et al., 2006), nos résultats sont en faveur de l'hypothèse selon laquelle une grande partie des étapes de planification conjointe puissent être réalisés par les mêmes processus utilisés lors des étapes de planification individuelle (voir : (Butterfill & Sinigaglia, 2012) Ainsi, à la différence du modèle classique *d'intentionnalité partagée* (Bratman, 1993) qui présuppose des représentations partagées des intentions des

deux agents, des modèles plus minimalistes de l'action conjointe font leur apparition (Vesper et al., 2010). Pour ces auteurs, les conditions minimales pour la réalisation d'une action conjointe seraient la possibilité que chaque agent puisse simplement planifier les actions de l'autre afin d'en prédire les conséquences. Aussi, une coordination pourrait s'établir bien que certaines actions planifiées puissent ne pas avoir d'agent associé (« *agent-neutral intentions* »). Autrement dit, il serait possible d'établir des plans à minima des actions à réaliser (par exemple dans quel ordre), préalablement au fait de sélectionner qui de l'un ou l'autre des agents va réaliser ces actions. Seul le « but » commun de la tâche serait partagé, et la coordination entre les deux agents s'établirait de la même manière s'il s'agissait d'un seul agent qui devait par exemple faire passer un objet de sa main droite à sa main gauche. Ainsi bien que les deux effecteurs possèdent des planifications séparées, le fait de prédire les conséquences de l'action du premier effecteur permet simplement d'estimer temporellement et spatialement les commandes du deuxième effecteur pour faire passer l'objet correctement d'une main à l'autre. Pour Stephen Butterfill, une coordination peut s'établir sans qu'il y ai besoin que les deux personnes utilisent des fonctions cognitives de haut niveau telles que la compréhension des états mentaux d'autrui (ou « théorie de l'esprit »). Toutefois, les nombreuses interactions entre les individus faciliteraient par ailleurs l'émergence de fonctions de cognition sociale plus sophistiquées chez l'homme (voir : (Butterfill, 2012). Enfin, pour d'autres auteurs, il est possible que des effets de coordination interindividuelle apparaissent bien que ni l'un ou l'autre des agents ne soit conscient d'être impliqué dans une action conjointe, insistant sur l'idée que certains aspects de plus bas niveau que la théorie de l'esprit peuvent être des éléments constitutifs de l'établissement des raisonnements collectifs, ou appelés aussi raisonnement de groupe (voir : (Pacherie, 2011).

En conclusion, nos résultats vont dans le sens d'un modèle des interactions sociales basé sur des propriétés de bas niveau conséquentes des anticipations sensori-motrices chez l'agent et utilisés en tant qu'indices non-verbaux chez l'observateur. Cependant nous n'excluons pas la possibilité d'indicateurs complémentaires pour aider à la compréhension des

actions. En conséquence, nous proposons replacer nos résultats dans les trois champs disciplinaires dans lesquels ces travaux se situent.

### 6.3. Perspectives en Neurosciences

Au niveau neuronal, s'il semble clair que les neurones miroirs puissent être à l'origine de la correspondance entre action et observation dans le cortex, de nouvelles propositions doivent être faites concernant le rôle des structures sous-corticales pour ce type de tâche et leurs relation avec les structures corticales. Par exemple, Caligiore, Pezzulo, Miall et Baldassarre (2013) proposent une description particulièrement intéressante des systèmes neuronaux impliquant le STS (jonction parétiotemporelle), l'aire PFG (cortex pariétal) et l'aire F5 (cortex prémoteur) pour y intégrer des structures sous-corticales comme le cervelet et les ganglions de la base, ainsi que le cortex préfrontal (Figure 1d).

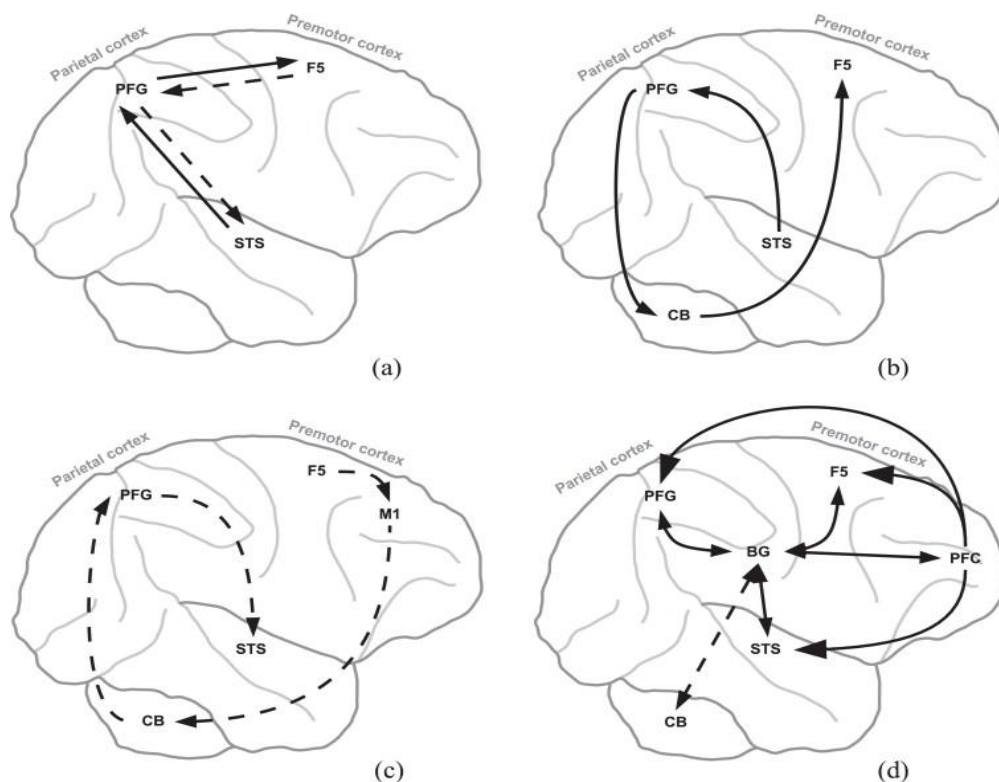


Figure 1. Modèles internes et circuits miroirs. Adapté de Caligiore, D., Pezzulo, G., Miall, R. C., & Baldassarre, G. (2013). The contribution of brain sub-cortical loops in the expression and acquisition of action understanding abilities. *Neuroscience & Biobehavioral Reviews*.

Dans leur modèle, les auteurs (Caligiore, Pezzulo, Miall, & Baldassarre, 2013) décrivent les circuits impliqués dans l'observation des action à travers le système miroir

formé par les aires STS (sillon temporal supérieur), l'aire (PFG) dans le lobe pariétal inférieur et le cortex prémoteur ventral (F5). Ils proposent que le circuit miroir fonctionne comme un *modèle inverse*, c'est-à-dire qu'il transforme la représentation visuelle de l'action codée par les neurones STS en un plan moteur codé par les neurones miroirs de l'aire F5 (Figure 1a, flèches pleines). Réciproquement, pendant l'exécution des actions, les connections de l'aire F5 vers le PFG et le STS pourrait fonctionner comme un *modèle direct*, qui converti le plan moteur en conséquences sensorielles de l'action. Il s'agirait d'une représentation visuelle prédite, codée par les neurones de l'aire STS (Figure 1a, flèches en pointillé). Cependant, les auteurs soulignent que le circuit miroir ne fonctionne pas de manière isolée des structures sous-corticales. Etant donné le rôle central du cervelet comme structure à l'origine du fonctionnement des modèles (Kawato, 1999; Miall, 2003; Wolpert et al., 1998), Miall (2003) propose comment le circuit miroir pourrait impliquer le cervelet (Figure 1. b,c). D'après les connections fortes entre les aires pariétales et le cervelet (voir : (Middleton & Strick, 2000) qui est lui-même connecté avec l'aire premotrice ventrale (Strick, Dum, & Fiez, 2009), il semblerait qu'il puisse fonctionner comme un *modèle inverse* (Figure 1b). Mais par ailleurs, le cervelet pourrait utiliser les copies efférentes des signaux moteurs du cortex moteur primaire (M1) pour fonctionner comme un *modèle direct* dont l'objectif serait générer une prédiction sensorielle des conséquences de l'action impliquant une mise à jour visuo-motrice dans l'aire PFG et une activation sensorielle dans l'aire STS (Figure 1c). Pour appuyer cette idée, Cattaneo (2012) propose un rôle important du cervelet dans le séquençage des action en montrant que des patients lésés sont moins bons que des sujets contrôles pour assembler des images d'une action en une séquence ordonnée (Cattaneo et al., 2012). Calgiore et collaborateurs (2013) vont encore plus loin en proposant que le cortex préfrontal, connu pour son rôle dans nos capacités à nous projeter dans le futur serait un prérequis important pour détecter des intentions ou des buts de haut niveau, considérés comme des états désirés anticipés. Ils soulignent le rôle du cortex préfrontal dans la sélection des actions reliées à ces buts à travers ses projections vers le circuit miroir. Les mécanismes de sélection des actions



notamment supportés par les boucles entre les ganglions de la base et le cortex pourrait s'articuler autour de l'hypothèse de la « compétition d'affordance » (Cisek & Kalaska, 2010; Cisek, 2007). Dans ce cadre, les circuits pariéto-frontaux proposent des plans d'actions multiples qui entrent en compétition jusqu'à ce qu'un soit sélectionné et exécuté. La voie dorsale est reliée à la spécification des actions et à la préparation motrice alors que la voie ventrale est reliée à l'identité des objets et aux processus affectifs et de motivation internes. Au niveau des ganglions de la base cela permet de générer des biais comportementaux qui participent à l'évaluation des plans moteurs au cours de leur formation (Figure 2).

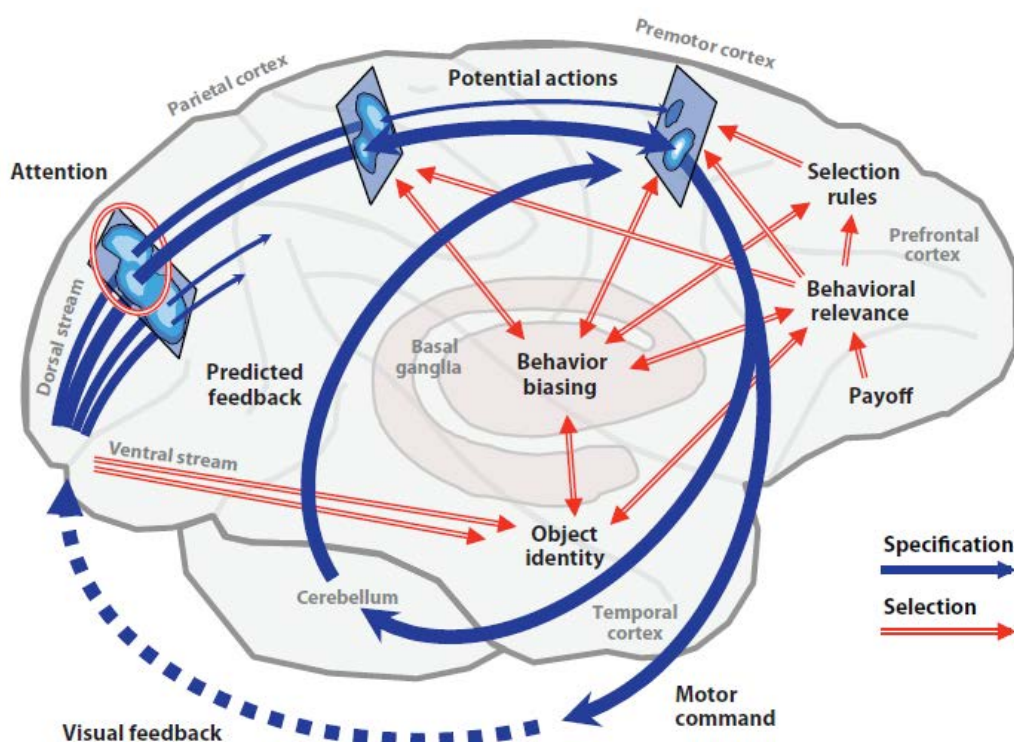


Figure 2. Schéma de l'hypothèse de la « compétition d'affordance » dans le cadre du mouvement visuellement guidé. Adapté de Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual review of neuroscience*, 33, 269–298.

Ainsi ces mécanismes de compétition lors de l'exécution présentés très brièvement ici permettent de mieux comprendre comment les actions peuvent être sélectionnées à partir de contrôles de plus haut niveau et comment des mécanismes de séquençage des actions peuvent s'articuler avec ces mécanismes de sélection afin de générer des prédictions lors de l'observation des actions.

#### 6.4. Perspectives en Psychologie Cognitive

Au niveau cognitif, il semblerait qu'il existe deux systèmes complémentaires pour la compréhension des actions (Centelles, Assaiante, Nazarian, Anton, & Schmitz, 2011; de Lange, Spronk, Willems, Toni, & Bekkering, 2008; Decety & Grèzes, 1999). Le premier système, dit de simulation motrice, effectue une correspondance directe entre l'action observée dans un format correspondant au répertoire moteur de l'observateur. Ce système permet par un modèle direct de déterminer le but d'une action considérant les commandes motrices observées. Cependant, dans le cas où les actions sont peu spécifiées, difficiles à observer ou impossibles à simuler dans le répertoire moteur humain, d'autres systèmes d'inférences peuvent prendre le relais (Csibra, 2008; Friston et al., 2011; Gergely & Csibra, 2003). L'idée est que ces systèmes ne soient pas forcément en compétition mais plutôt qu'ils viennent renseigner sur des aspects différents de l'action (Stapel et al., 2012) et leur mise en jeu pourrait dépendre à la fois de la qualité de la source cinématique des informations, des contraintes de la tâche ou du contexte des actions ((Van Overwalle & Baetens, 2009). Par exemple, dans notre expérience en observation, si la première partie du mouvement peut-être simulée directement, un modèle direct pourrait s'avérer parfois insuffisant pour engendrer une réponse correcte, étant donné que la variabilité de la cinématique des mouvements contient une grande part de chevauchement entre les catégories. Il est possible qu'un système d'inférence complémentaire, puisse venir aider à décrypter ces informations en intégrant les information cinématiques avec les informations contextuelles pour aider à décoder les intentions du mouvement. Par exemple Chambon et collaborateurs (2011) ont montré des interactions entre les attentes préalables (prior expectations) et les informations perceptuelles avec des modulations de ces interactions en fonction du type d'intentions (basique ou superordonnée ou sociale vs non-sociale). Par exemple, lorsque la qualité de l'information perceptuelle diminuait, l'effet de biais (la contribution des attentes préalables) était accru avec des réponses des participants pour des intentions comme étant les causes les plus probables de l'action observée (Chambon et al., 2011). Ainsi les auteurs proposent que les jugements

intentionnels sont basés sur un équilibre relatif entre les informations sensorielles (bottom-up) et les attentes préalables (top-down), et plaident en faveur d'un modèle hybride de la compréhension des actions.

Le cadre théorique du codage prédictif proposé récemment (Clark, 2013; Friston et al., 2011; Kilner et al., 2007) semble être en mesure d'expliquer ces résultats. Ce cadre théorique repose sur l'idée d'une architecture hiérarchique concernant la représentation des actions du niveau intentionnel au niveau cinématique (Hamilton & Grafton, 2007) notamment observé en IRMf chez l'humain (Fuster, 2004; Koechlin, Ody, & Kouneiher, 2003). L'idée du codage prédictif est que chaque niveau supérieur de la hiérarchie agit comme une contrainte préalable (priors) qui chercherait à prédire les entrées des niveaux inférieurs sur la base de leur modèle d'émergence et de la structure causale du monde (Clark, 2013). Les erreurs dans la prédiction des niveaux inférieurs entraîneraient les niveaux supérieurs à s'adapter pour réduire cet écart. Ce processus, qui opère à des niveaux multiples, permettrait alors au cerveau d'encoder un contenu informationnel important à propos de la source des signaux qui le perturbe régulièrement. Ainsi, ce type de codage (prédictif) cherche à améliorer les standards computationnels tels que les apprentissages basés sur des algorithmes de rétro-propagation. En utilisant les connections top-down pour obtenir des états désirés des unités des couches cachées, le système va superviser lui-même le développement de ses modèles de reconnaissance en utilisant des modèles *génératifs* qui cherchent à réaliser des patterns sensoriels pour lui-même. Un modèle génératif cherche à capturer les structures statistiques d'un set d'entrée en « récapitulant » les causes responsables de cette structure. Pour la vision, un modèle génératif cherchera à capturer comment les réponses des systèmes sensoriels de bas niveau sont générés par les aspects visuels présentés dans la scène. En pratique, cela induit que les connections top-down vont encoder un modèle probabiliste des activités des groupes d'unités des niveaux inférieurs mais surtout que l'information propagée vers les niveaux supérieurs sont uniquement les variations « inattendues » c'est-à-dire *l'erreur de prédiction* ou la différence entre le signal réel et le signal prédit. Ce type de codage permet

d'éviter un grand nombre de redondance et de permettre une forte libération de ressources cognitives. Clark (2013) précise que pour les processus prédictifs des actions, les « intentions » extraient activement le résultat des flux sensoriels attendus que le cerveau prédit à travers le déploiement de ces intentions sous forme d'éléments moteurs détaillés. Cette hypothèse trouverait un écho intéressant dans le cas de l'*inférence active*, lorsque l'agent bouge ses senseurs de manière à maximiser la quantité d'informations sensorielles attendues en conséquence des actions exécutées ou perçues (Friston et al., 2011; Friston, 2009, 2010). Pour Andy Clark (2013), ce cadre théorique élégant mêle à la fois perception, action et cognition « qui agissent de concert pour minimiser les erreurs de prédiction sensorielle en échantillonnant sélectivement et en sculptant activement la matrice du stimulus. Ils conspirent ensemble pour mouvoir la créature dans le temps et l'espace de manière à satisfaire des attentes (...) changeant constamment. ».

Pour conclure, au niveau cognitif, plusieurs théories sont encore largement en discussion (voir : (Zentgraf, Munzert, Bischoff, & Newman-Norlund, 2011). Nos résultats suggèrent clairement une implication des informations de bas-niveau qui pourraient jouer un rôle clé dans l'émergence et l'apparition de fonction cognitives plus avancées. Cependant, de nouvelles expériences sont nécessaires pour mieux comprendre l'articulation entre les informations de haut niveau et de bas-niveau. Les modèles hybrides basés sur la complémentarité des différentes sources d'informations semblent être des pistes intéressantes pour expliquer l'étendue des capacités de compréhension des actions chez l'homme.

## **6.5. Perspectives en modèles computationnels pour la Robotique.**

Enfin au niveau robotique, une des questions importantes est de déterminer comment les systèmes de compréhension des actions pourraient se *développer* ou émerger tels que dans les circuits miroirs humains par exemple. L'apprentissage associatif séquentiel (*associative*

*sequence learning* : ASL) est une hypothèse qui propose que le problème de la correspondance directe entre perception et action au niveau neuronal ou les comportements d'imitation (Catmur, Walsh, & Heyes, 2009) soient en réalité le résultat de l'expérience sensori-motrice acquise pendant le développement. Un des exemples en robotique est l'idée développée par l'équipe de Nagai et collaborateurs selon laquelle des neurones formels aux propriétés miroirs peuvent émerger de l'aspect immature de la vision dans les premiers âges du développement (voir : (Baraglia, Nagai, Kawai, & Asada, 2012; Nagai, Kawai, & Asada, 2011). D'autres modèles basés sur la théorie des champs de neurones dynamiques (Dynamic Field Theory, voir (Erlhagen & Schöner, 2002; Schneegans & Schöner, 2008; Schöner, 2008) sont particulièrement intéressants compte-tenu de leur pouvoir de représentation de l'information sous la forme de champs d'activation dynamiques qui traduisent en temps réel le monde perçu par un système incarné (par ex : robotique). Un exemple de l'utilisation de ces modèles pour la robotique humanoïde sont les recherches menées par le groupe d'Estella Bicho à l'université de Minho sur les actions conjointes avec un robot humanoïde (voir : (Bicho et al., 2012; Bicho, Erlhagen, Louro, & Costa e Silva, 2011; Bicho, Louro, & Erlhagen, 2010). Ainsi, dans le cadre de nos travaux collaboratifs au sein de l'ANR « INTERACT » (2009-2013), les chercheurs de l'Etis ont pu proposer une architecture computationnelle, reprenant à la fois les résultats obtenus empiriquement en cinématique et des structures sous-corticales impliqués dans le contrôle moteur : le cortex moteur primaire, l'hippocampe, le cervelet et le striatum (Bailly, Andry, & Gaussier, 2012). Ce type d'architecture peut notamment être utilisé pour l'apprentissage par démonstration de tâches séquentielles (de Rengervé et al., 2011) mais aussi de permettre des apprentissages via des signaux externes donné par un humain (De Rengervé, Braud, Andry, & Gaussier, 2012) ou alors en utilisant les rythmes des interactions comme signal de renforcement et d'apprentissage (Andry, Blanchard, & Gaussier, 2011). Ces dernières propositions sont particulièrement enthousiasmantes en psychologie cognitive car elles proposent de nouveaux

questionnements sur les interactions sociales humaines et leurs émergences spontanées dans le cas d'interactions notamment non-intentionnelles entre les participants.

Pour résumer, les modèles computationnels classiques utilisés pour le contrôle moteur en robotique humanoïde bien que adaptés pour décrire certains comportements biologiques semblent être insuffisants pour décrire l'étendue complète des informations prises en compte lors de la planification des actions, en particulier dans le cadre des interactions sociales. Etant donné l'attraction qu'un dispositif robotique engendre en particulier lorsque son aspect esthétique est proche de celui d'un humain, les perspectives en réhabilitation motrice ou en remédiation cognitive sont très importantes. Cependant, ces perspectives semblent être concomitantes avec les capacités d'interaction « intuitives » du robot. Autrement dit, il semble important de faire avancer les modèles de contrôle moteur des robots humanoïdes pour éviter de tomber dans les écueils de la « uncanny valley » qui génèrent des erreurs de prédiction importantes chez les partenaires du robot. Par ailleurs, le fait de nécessiter de nouveaux apprentissages spécifiques pour interagir avec le système pourraient entraîner des découragements, comme c'est le cas pour des certains dispositifs artificiels conçus pourtant pour être interactifs (coachs virtuels, écrans vidéo, montres « intelligentes », etc...).

Enfin, les questionnements apportés par la création de systèmes incarnés et leur mise en relation directe avec leur environnement, dans le cas des robots humanoïdes, sont particulièrement pertinents pour mieux comprendre certaines capacités cérébrales complexes à la fois chez l'adulte et au cours du développement chez l'enfant. Ces nouvelles pistes de réflexion constituent à la fois des angles de lecture nouveaux et permettent l'émergence de nouvelles problématiques qui auraient pu être négligés en neurosciences et en psychologie. Par ailleurs, les architectures robotiques permettent une mise à l'épreuve directe des modèles cognitifs permettant ainsi des aller-retour rapides et importants entre théorie et expérimentation. C'est pourquoi nous souhaitons pour conclure insister sur l'intérêt du caractère interdisciplinaire des recherches menées en robotique en soulignant une nouvelle

fois l'apport des données comportementales humaines pour le développement de modèles innovants.

# Références

- Adam, J. J., Nieuwenstein, J. H., Huys, R., Paas, F. G. W. C., Kingma, H., Willems, P., & Werry, M. (2000). Control of rapid aimed hand movements: The one-target advantage. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 295.
- Adam, J. J., Paas, F. G. W. C., Eyssen, I. C. J. M., Slingerland, H., Bekkering, H., & Drost, M. (1995). The control of two-element, reciprocal aiming movements: Evidence for chunking. *Human Movement Science*, 14, 1-11.
- Adam, J. J., van der Bruggen, D. P. W., & Bekkering, H. (1993). The control of discrete and reciprocal target-aiming responses: Evidence for the exploitation of mechanics. *Human Movement Science*, 12, 353-364.
- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature neuroscience*, 11, 1109–1116.
- Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of neurophysiology*, 106, 1437–1442.
- Andry, P., Blanchard, A., & Gaussier, P. (2011). Using the rhythm of nonverbal human–robot interaction as a signal for learning. *Autonomous Mental Development, IEEE Transactions on*, 3, 30–42.
- Andry, P., Gaussier, P., Moga, S., Banquet, J. P., & Nadel, J. (2001). Learning and communication via imitation: An autonomous robot perspective. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 31, 431–442.
- Anquetil, T., & Jeannerod, M. (2007). Simulated actions in the first and in the third person perspectives share common representations. *Brain research*, 1130, 125–129.
- Ansuini, C., Giosa, L., Turella, L., Altoè, G., & Castiello, U. (2008). An object for an action, the same object for other actions: effects on hand shaping. *Experimental Brain Research*, 185, 111-119.
- Ansuini, C., Grigis, K., Adamovich, S., & Castiello, U. (2009). Breaking the flow of an



- action. *Experimental Brain Research*, 192, 287-292.
- Ansuini, C., Santello, M., Massaccesi, S., & Castiello, U. (2006). Effects of end-goal on hand shaping. *Journal of neurophysiology*, 95, 2456–2465.
- Armbrüster, C., & Spijkers, W. (2006). Movement planning in prehension: do intended actions influence the initial reach and grasp movement? *Motor control*, 10, 311-329.
- Athenes, S. (1992). *Aspects of motor coordination: A study of prehension and handwriting posture* (Phd Thesis). Université d'Aix-Marseille 2, France.
- Bailly, D., Andry, P., & Gaussier, P. (2012). Learning anticipatory motor control. In *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)* (p. 1-2). Présenté à 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL).
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., ... Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, 30, 917-926.
- Baraglia, J., Nagai, Y., Kawai, Y., & Asada, M. (2012). The role of temporal variance in motions for the emergence of mirror neurons systems. In *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on* (p. 1–2).
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Frontiers in Human Neuroscience*, 6. doi:10.3389/fnhum.2012.00117
- Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008a). The case of Dr. Jekyll and Mr. Hyde: A kinematic study on social intention. *Consciousness and Cognition*, 17, 557-564.
- Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008b). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition*, 106, 894-912.
- Becchio, C., Sartori, L., & Castiello, U. (2010). Toward You The Social Side of Actions.

*Current Directions in Psychological Science*, 19, 183-188.

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science*, 13, 370–374.

Belmalih, A., Borra, E., Contini, M., Gerbella, M., Rozzi, S., & Luppino, G. (2007). A multiarchitectonic approach for the definition of functionally distinct areas and domains in the monkey frontal lobe. *Journal of Anatomy*, 211, 199–211.

Bernstein, N. A. (1967). The co-ordination and regulation of movements.

Bicho, E., Erlhagen, W., Louro, L., & Costa e Silva, E. (2011). Neuro-cognitive mechanisms of decision making in joint action: A human–robot interaction study. *Human movement science*, 30, 846–868.

Bicho, E., Erlhagen, W., Sousa, E., Louro, L., Hipolito, N., Silva, E. C., ... Hulstijn, M. (2012). The power of prediction: Robots that read intentions. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (p. 5458–5459).

Bicho, E., Louro, L., & Erlhagen, W. (2010). Integrating verbal and nonverbal communication in a dynamic neural field architecture for human–robot interaction. *Frontiers in neurorobotics*, 4.

Blythe, P. W., Todd, P. M., & Miller, G. F. (1999). How motion reveals intention: Categorizing social interactions. In *Simple heuristics that make us smart* (p. 257-285). New York, NY, US: Oxford University Press.

Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F., & Fogassi, L. (2010). Ventral Premotor and Inferior Parietal Cortices Make Distinct Contribution to Action Organization and Intention Understanding. *Cerebral Cortex*, 20, 1372 -1385.

Bootsma, R., Marteniuk, R., MacKenzie, C., & Zaal, F. J. M. (1994). The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Experimental Brain Research*, 98. doi:10.1007/BF00233990

Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., & Luppino, G.

- (2008). Cortical Connections of the Macaque Anterior Intraparietal (AIP) Area. *Cerebral Cortex*, 18, 1094 -1111.
- Boucher, J.-D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., ... Ventre-Dominey, J. (2012). I reach faster when I see you look: gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in neurorobotics*, 6.
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta psychologica*, 106, 3–22.
- Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and cognition*, 44, 124–143.
- Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current Biology*, 17, 2117–2121.
- Bratman, M. E. (1993). Shared Intention. *Ethics*, 104, 97-113.
- Breazeal, C., Buchsbaum, D., Gray, J., Gatenby, D., & Blumberg, B. (2005). Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life*, 11, 31–62.
- Breazeal, C., Gray, J., & Berlin, M. (2009). An embodied cognition approach to mindreading skills for socially intelligent robots. *The International Journal of Robotics Research*, 28, 656–680.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13, 400–404.
- Buneo, C. A., & Andersen, R. A. (2006). The posterior parietal cortex: Sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia*, 44, 2594-2606.
- Butterfill, S. A. (2012). Interacting mindreaders. *Philosophical Studies*, 165, 841-863.

- Butterfill, S. A., & Sinigaglia, C. (2012). Intention and motor representation in purposive action. *Philosophy and Phenomenological Research*.
- Caligiore, D., Pezzulo, G., Miall, R. C., & Baldassarre, G. (2013). The contribution of brain sub-cortical loops in the expression and acquisition of action understanding abilities. *Neuroscience & Biobehavioral Reviews*. doi:10.1016/j.neubiorev.2013.07.016
- Campione, G. C., & Gentilucci, M. (2011). Is the observation of the human kinematics sufficient to activate automatic imitation of transitive actions? *Behavioural Brain Research*, 225, 201-208.
- Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, 100, 5497-5502.
- Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, 125, 1839-1849.
- Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 2369-2380.
- Cattaneo, L., Fasanelli, M., Andreatta, O., Bonifati, D. M., Barchiesi, G., & Caruana, F. (2012). Your actions in my cerebellum: subclinical deficits in action observation in patients with unilateral chronic cerebellar stroke. *Cerebellum (London, England)*, 11, 264-271.
- Centelles, L., Assaiante, C., Nazarian, B., Anton, J.-L., & Schmitz, C. (2011). Recruitment of Both the Mirror and the Mentalizing Networks When Observing Social Interactions Depicted by Point-Lights: A Neuroimaging Study. *PLoS ONE*, 6, e15749.
- Chambon, V., Domenech, P., Pacherie, E., Koechlin, E., Baraduc, P., & Farrer, C. (2011). What Are They Up To? The Role of Sensory Evidence and Prior Knowledge in Action Understanding. *PLoS ONE*, 6, e17133.

- Chaminade, T., & Cheng, G. (2009). Social cognitive neuroscience and humanoid robotics. *Journal of Physiology-Paris*, 103, 286–295.
- Chaminade, T., Franklin, D. W., Oztop, E., & Cheng, G. (2005). Motor interference between Humans and Humanoid Robots: Effect of Biological and Artificial Motion. In *The 4th International Conference on Development and Learning, 2005. Proceedings* (p. 96-101). Présenté à The 4th International Conference on Development and Learning, 2005. Proceedings, IEEE.
- Chieffi, S., & Gentilucci, M. (1993). Coordination between the transport and the grasp components during prehension movements. *Experimental Brain Research*, 94, 471–477.
- Chong, T. T.-J., Cunnington, R., Williams, M. A., & Mattingley, J. B. (2009). The role of selective attention in matching observed and executed actions. *Neuropsychologia*, 47, 786-795.
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 1585–1599.
- Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual review of neuroscience*, 33, 269–298.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181-204.
- Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., & Martineau, J. (1998). Perception of motion and qEEG activity in human adults. *Electroencephalography and Clinical Neurophysiology*, 107, 287-295.
- Cohen, R. G., & Rosenbaum, D. A. (2004). Where grasps are made reveals how grasps are planned: generation and recall of motor plans. *Experimental Brain Research*, 157, 486–495.
- Cook, J., Saygin, A. P., Swain, R., & Blakemore, S.-J. (2009). Reduced sensitivity to

- minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia*, 47, 3275–3278.
- Costantini, M., Ambrosini, E., Scorolli, C., & Borghi, A. M. (2011). When objects are close to me: affordances in the peripersonal space. *Psychonomic bulletin & review*, 18, 302–308.
- Costantini, M., Committeri, G., & Sinigaglia, C. (2011). Ready both to your and to my hands: mapping the action space of others. *PloS one*, 6, e17923.
- Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia*, 40, 492–502.
- Crawford, J. D., Medendorp, W. P., & Marotta, J. J. (2004). Spatial transformations for eye–hand coordination. *Journal of neurophysiology*, 92, 10–19.
- Cross, E. S., Kraemer, D. J. M., Hamilton, A. F. de C., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the Action Observation Network to Physical and Observational Learning. *Cerebral Cortex*, 19, 315–326.
- Csibra, G. (2008). Action mirroring and action understanding: An alternative account. *Sensorymotor Foundations of Higher Cognition. Attention and Performance XXII*, 435–459.
- Davidson, P. R., & Wolpert, D. M. (2004). Internal models underlying grasp can be additively combined. *Experimental Brain Research*, 155, 334–340.
- Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: a short review. *Journal of Neural Engineering*, 2, S313.
- Davy, D. T., & Audu, M. L. (1987). A dynamic optimization technique for predicting muscle forces in the swing phase of gait. *Journal of Biomechanics*, 20, 187–201.
- De Lange, F. P., Spronk, M., Willems, R. M., Toni, I., & Bekkering, H. (2008). Complementary Systems for Understanding Action Intentions. *Current Biology*, 18, 454–457.
- De Rengervé, A., Braud, R., Andry, P., & Gaussier, P. (2012). Behavior adaptation from

- negative social feedback based on goal awareness. In *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on* (p. 1–6).
- De Rengervé, A., Hirel, J., Andry, P., Quoy, M., & Gaussier, P. (2011). On-line learning and planning in a pick-and-place task demonstrated through body manipulation. In *2011 IEEE International Conference on Development and Learning (ICDL)* (Vol. 2, p. 1-6). Présenté à 2011 IEEE International Conference on Development and Learning (ICDL), IEEE.
- Decety, J. (1993). Analysis of actual and mental movement times in graphic tasks. *Acta psychologica*, 82, 367–372.
- Decety, J., & Chaminade, T. (2003a). Neural correlates of feeling sympathy. *Neuropsychologia*, 41, 127–138.
- Decety, J., & Chaminade, T. (2003b). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and cognition*, 12, 577–596.
- Decety, J., Chaminade, T., Grèzes, J., & Meltzoff, A. N. (2002). A PET Exploration of the Neural Mechanisms Involved in Reciprocal Imitation. *NeuroImage*, 15, 265-272.
- Decety, J., & Grèzes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in cognitive sciences*, 3, 172–178.
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and cognitive neuroscience reviews*, 3, 71–100.
- Decety, J., & Jeannerod, M. (1995). Mentally simulated movements in virtual reality: does Fitt's law hold in motor imagery? *Behavioural Brain Research*, 72, 127–134.
- Decety, J., & Michel, F. (1989). Comparative analysis of actual and mental movement times in two graphic tasks. *Brain and cognition*, 11, 87–97.
- Dechent, P., Merboldt, K.-D., & Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cognitive Brain Research*.
- Delevoeye-Turrell, Y., Giersch, A., Wing, A. M., & Danion, J.-M. (2007). Motor fluency

- deficits in the sequencing of actions in schizophrenia. *Journal of abnormal psychology*, 116, 56.
- Delevoeye-Turrell, Y., & Wing, A. M. (2004). Action and motor skills: adaptive behaviour for intended goals. *Handbook of Cognition*, 130–157.
- Desmurget, M., & Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences*, 13, 411-419.
- Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, 10, 732-739.
- Eastough, D., & Edwards, M. G. (2007). Movement kinematics in prehension are affected by grasping objects of different mass. *Experimental brain research*, 176, 193–198.
- Ebisch, S. J. H., Perrucci, M. G., Ferretti, A., Del Gratta, C., Romani, G. L., & Gallese, V. (2008). The Sense of Touch: Embodied Simulation in a Visuotactile Mirroring Mechanism for Observed Animate or Inanimate Touch. *Journal of Cognitive Neuroscience*, 20, 1611-1623.
- Engel, K. C., Flanders, M., & Soechting, J. F. (1997). Anticipatory and sequential motor control in piano playing. *Experimental brain research*, 113, 189–199.
- Erlhagen, W., & Schöner, G. (2002). Dynamic Field Theory of Movement Preparation. *Psychological Review*, 109, 545–572.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, 15, 399–402.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of neurophysiology*, 73, 2608–2611.
- Feil-Seifer, D., & Mataric, M. J. (2005). Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on* (p. 465–468).
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., & Gentilucci, M. (2011). Social



- Requests and Social Affordances: How They Affect the Kinematics of Motor Sequences during Interactions between Conspecifics. *PLoS ONE*, 6, e15855.
- Fikes, T. G., Klatzky, R. L., & Lederman, S. J. (1994). Effects of object texture on precontact movement time in human prehension. *Journal of motor behavior*, 26, 325–332.
- Fischer, M. H., Rosenbaum, D. A., & Vaughan, J. (1997). Speed and sequential effects in reaching. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 404-428.
- Fischman, M. G., & Reeve, T. G. (1992). Slower movement times may not necessarily imply online programming. *Journal of Human Movement Studies*, 22, 131–144.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology; Journal of Experimental Psychology*, 47, 381-391.
- Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, 424, 769–771.
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *The journal of Neuroscience*, 5, 1688–1703.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal Lobe: From Action Organization to Intention Understanding. *Science*, 308, 662 -667.
- Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, 42, 143–166.
- Frak, V., Paulignan, Y., & Jeannerod, M. (2001). Orientation of the opposition axis in mentally simulated grasping. *Experimental Brain Research*, 136, 120–127.
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences*, 13, 293–301.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11, 127-138.
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference.

*Biological cybernetics*, 104, 137-160.

Furlanetto, T., Cavallo, A., Manera, V., Tversky, B., & Becchio, C. (2013). Through your eyes: incongruence of gaze and action increases spontaneous perspective taking. *Frontiers in human neuroscience*, 7.

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in cognitive sciences*, 8, 143–145.

Fuster, J. M. (2008). *The prefrontal cortex*. London: Academic Press.

Gallese, V. (2001). The shared manifold hypothesis. From mirror neurons to empathy. *Journal of consciousness studies*, 8, 5–7.

Gallese, V. (2005). Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, 4, 23-48.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593 -609.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In *Common mechanisms in perception and action: Attention and Performance XIX* (p. 247-266). Oxford University Press.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, 2, 493–501.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8, 396-403.

Gaussier, P., Moga, S., Quoy, M., & Banquet, J. P. (1998). From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence*, 12, 701–727.

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, 19, 1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain:

- The mirror neuron system responds to human and robotic actions. *NeuroImage*, 35, 1674-1684.
- Gentilucci, M., Castiello, U., Corradini, M. L., Scarpa, M., Umiltà, C., & Rizzolatti, G. (1991). Influence of different types of grasping on the transport component of prehension movements. *Neuropsychologia*, 29, 361–378.
- Gentilucci, M., Negrotti, A., & Gangitano, M. (1997). Planning an action. *Experimental Brain Research*, 115, 116-128.
- Georgiou, I., Becchio, C., Glover, S., & Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition*, 102, 415-433.
- Gerardin, E., Sirigu, A., Lehericy, S., Poline, J.-B., Gaymard, B., Marsault, C., ... Bihan, D. L. (2000). Partially Overlapping Neural Networks for Real and Imagined Hand Movements. *Cerebral Cortex*, 10, 1093-1104.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in cognitive sciences*, 7, 287–292.
- Gianelli, C., Lugli, L., Baroni, G., Nicoletti, R., & Borghi, A. M. (2011). “The object is wonderful or prickly”: how different object properties modulate behavior in a joint context. *European perspectives on cognitive science*.
- Gianelli, C., Scorolli, C., & Borghi, A. M. (2013). Acting in perspective: the role of body and language as social tools. *Psychological research*, 77, 40–52.
- Gibson, J. J. (1977). The concept of affordances. *Perceiving, acting, and knowing*, 67–82.
- Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4, 179–192.
- Goodrich, M. A., & Schultz, A. C. (2007). Human-robot interaction: a survey. *Foundations and Trends in Human-Computer Interaction*, 1, 203–275.
- Gordon, J., Ghilardi, M. F., & Ghez, C. (1994). Accuracy of planar reaching movements. I. Independence of direction and extend variability. *Experimental Brain Research*, 99, 97–111.

- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *Neuroimage*, 36, T22–T32.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. *Experimental Brain Research*, 112. doi:10.1007/BF00227183
- Grèzes, J., Armony, J. ., Rowe, J., & Passingham, R. . (2003). Activations related to « mirror » and « canonical » neurones in the human brain: an fMRI study. *NeuroImage*, 18, 928-937.
- Grezes, J., & Decety, J. (2000). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human brain mapping*, 12, 1–19.
- Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does Perception of Biological Motion Rely on Specific Brain Regions? *NeuroImage*, 13, 775–785.
- Grosjean, M., Shiffrar, M., & Knoblich, G. (2007). Fitts's Law Holds for Action Perception. *Psychological Science*, 18, 95-99.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of cognitive neuroscience*, 12, 711–720.
- Guiard, Y. (1993). On Fitts's and Hooke's laws: Simple harmonic movement in upper-limb cyclical aiming. *Acta Psychologica*, 82, 139-159.
- Guiard, Y. (1997). Fitts' law in the discrete vs. cyclical paradigm. *Human Movement Science*, 16, 97-131.
- Guigon, E., Baraduc, P., & Desmurget, M. (2003). Constant effort computation as a determinant of motor behavior. *Advances in Computational Motor Control*, 2.
- Guigon, E., Baraduc, P., & Desmurget, M. (2008). Computational motor control: feedback and accuracy. *European Journal of Neuroscience*, 27, 1003–1016.

- Haggard, P. (1998). Planning of action sequences. *Acta Psychologica*, 99, 201-215.
- Hallett, M., Fieldman, J., Cohen, L. G., Sadato, N., & Pascual-Leone, A. (1994). Involvement of primary motor cortex in motor imagery and mental practice. *Behavioral and Brain Sciences*, 17, 210–210.
- Hamilton, A. F., & Grafton, S. T. (2007). The motor hierarchy: from kinematics to goals and intentions. *Sensorimotor foundations of higher cognition*, 22, 381–408.
- Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Van Gelderen, P., & Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *Journal of Neurophysiology*, 89, 989–1002.
- Hanson, D., Olney, A., Prilliman, S., Mathews, E., Zielke, M., Hammons, D., ... Stephanou, H. (2005). Upending the uncanny valley. In *Proceedings of the national conference on artificial intelligence* (Vol. 20, p. 1728).
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences*, 95, 15061 - 15065.
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394, 780-784.
- Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural computation*, 13, 2201–2220.
- Haruno, M., Wolpert, D. M., & Kawato, M. (2003). Hierarchical MOSAIC for movement generation. In *International congress series* (Vol. 1250, p. 575–590).
- Hasnain, S. K., Gaussier, P., & Mostafaoui, G. (2012). Synchrony as a tool to establish focus of attention for autonomous robots. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (p. 2423–2428).
- Hayes, A. E., Paul, M. A., Beuger, B., & Tipper, S. P. (2008). Self produced and observed actions influence emotion: The roles of action fluency and eye gaze. *Psychological*

*research*, 72, 461–472.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9, 188-194.

Hayhoe, M., Shrivastava, A., Mruczek, R., & Pelz, J. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, 3. doi:10.1167/3.1.6

Herbort, O., & Butz, M. V. (2007). Encoding Complete Body Models Enables Task Dependent Optimal Behavior (p. 1639-1644). Présenté à International Joint Conference on Neural Networks, 2007. IJCNN 2007, IEEE.

Herbort, Oliver, & Butz, M. V. (2009). Anticipatory Planning of Sequential Hand and Finger Movements. *Journal of Motor Behavior*, 41, 561-569.

Herbort, Oliver, & Butz, M. V. (2011). The continuous end-state comfort effect: weighted integration of multiple biases. *Psychological Research*, 76, 345-363.

Hesse, C., & Deubel, H. (2010). Advance Planning in Sequential Pick-and-Place Tasks. *Journal of Neurophysiology*, 104, 508-516.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6, 242-247.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of cognitive neuroscience*, 21, 1229–1243.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in cognitive sciences*, 8, 494–500.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and brain sciences*, 24, 849–878.

Horan, W. P., Kern, R. S., Shokat-Fadai, K., Sergi, M. J., Wynn, J. K., & Green, M. F. (2009). Social cognitive skills training in schizophrenia: An initial efficacy study of stabilized outpatients. *Schizophrenia Research*, 107, 47-54.

Iacoboni, M. (2009). Imitation, Empathy, and Mirror Neurons. *Annual Review of Psychology*,

- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the Intentions of Others with One's Own Mirror Neuron System. *PLoS Biology*, 3, e79.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical Mechanisms of Human Imitation. *Science*, 286, 2526-2528.
- Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, 9, 21-25.
- Jakobson, L. S., & Goodale, M. A. (1991). Factors affecting higher-order movement planning: a kinematic analysis of human prehension. *Experimental Brain Research*, 86. doi:10.1007/BF00231054
- Jeannerod, M. (1984). The timing of natural prehension movements. *Journal of Motor Behavior*, 16, 235-254.
- Jeannerod, M. (1988). *The neural and behavioural organization of goal-directed movements*. New York, NY, US: Clarendon Press/Oxford University Press.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, 17, 187–202.
- Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, 18, 314-320.
- Jeannerod, Marc. (2001). Neural Simulation of Action: A Unifying Mechanism for Motor Cognition. *NeuroImage*, 14, S103-S109.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, 14, 201–211.
- Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, 38, 379-393.
- Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–hand

- coordination in object manipulation. *the Journal of Neuroscience*, 21, 6917–6932.
- Johnson, S. H. (2000). Imagining the impossible: intact motor representations in hemiplegics. *Neuroreport*, 11, 729–732.
- Johnson-Frey, S., McCarty, M., & Keen, R. (2004). Reaching beyond spatial perception: Effects of intended future actions on visually guided prehension. *Visual Cognition*, 11, 371.
- Jordan, M. I. (1996). Computational aspects of motor control and motor learning. *Handbook of perception and action*, 2, 71–120.
- Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Social Neuroscience*, 4, 135–152.
- Kandel, S., Orliaguet, J.-P., & Boe, L.-J. (2000). Detecting anticipatory events in handwriting movements. *PERCEPTION-LONDON-*, 29, 953–964.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current opinion in neurobiology*, 9, 718–727.
- Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., & Yoshioka, T. (2003). Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. *Progress in brain research*, 142, 171–188.
- Ketcham, C. J., Seidler, R. D., Van Gemmert, A. W., & Stelmach, G. E. (2002). Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57, P54–P64.
- Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews Neuroscience*, 11, 417–428.
- Kilner, J. ., Paulignan, Y., & Blakemore, S. . (2003). An Interference Effect of Observed Biological Movement on Action. *Current Biology*, 13, 522–525.
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror



- neuron system. *Cognitive Processing*, 8, 159-166.
- Kilner, J. M., & Frith, C. D. (2008). Action observation: inferring intentions without mirror neurons. *Current Biology*, 18, R32–R33.
- Kilner, J. M., Marchant, J. L., & Frith, C. D. (2009). Relationship between Activity in Human Primary Motor Cortex during Action Observation and the Mirror Neuron System. *PLoS ONE*, 4, e4925.
- Kimberley, T. J., Khandekar, G., Skraba, L. L., Spencer, J. A., Van Gorp, E. A., & Walker, S. R. (2006). Neural substrates for motor imagery in severe hemiparesis. *Neurorehabilitation and Neural Repair*, 20, 268–277.
- Klatzky, R. L., Fikes, T. G., & Pellegrino, J. W. (1995). Planning for hand shape and arm transport when reaching for objects. *Acta Psychologica*, 88, 209-232.
- Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, 15, 99–104.
- Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 2021-2031.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302, 1181–1185.
- Koelkebeck, K., Pedersen, A., Suslow, T., Kueppers, K. A., Arolt, V., & Ohrmann, P. (2010). Theory of Mind in first-episode schizophrenia patients: Correlations with cognition and personality traits. *Schizophrenia Research*, 119, 115-123.
- Lackner, J. R., & Dizio, P. (1994). Rapid adaptation to Coriolis force perturbations of arm trajectory. *Journal of neurophysiology*, 72, 299–313.
- Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta psychologica*, 54, 115–130.
- Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London. Series B: Biological*

- Sciences*, 352, 1231–1239.
- Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature neuroscience*, 3, 1340–1345.
- Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *PERCEPTION-LONDON-*, 28, 1311–1328.
- Lewkowicz, D., Delevoye-Turrell, Y., Bailly, D., Andry, P., & Gaussier, P. (2013). Reading motor intention through mental imagery. *Adaptive Behavior*. doi:10.1177/1059712313501347
- Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences*, 106, 9925–9930.
- Lotze, M., & Cohen, L. G. (2006). Volition and imagery in neurorehabilitation. *Cognitive and behavioral neurology*, 19, 135–140.
- Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., ... Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *Journal of cognitive neuroscience*, 11, 491–501.
- MacKenzie, C. L., Marteniuk, R. G., Dugas, C., Liske, D., & Eickmeier, B. (1987). Three-dimensional movement trajectories in Fitts' task: Implications for control. *The Quarterly Journal of Experimental Psychology*, 39, 629–647.
- Macneilage, P. F. (1980). Speech Production. *Language and Speech*, 23, 3-23.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102, 59–70.
- Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research*, 211, 547-556.
- Manera, V., Schouten, B., Becchio, C., Bara, B. G., & Verfaillie, K. (2010). Inferring

- intentions from biological motion: A stimulus set of point-light communicative interactions. *Behavior research methods*, 42, 168–178.
- Martel, L., Bidet-Ildei, C., & Coello, Y. (2011). Anticipating the terminal position of an observed action: Effect of kinematic, structural, and identity information. *Visual Cognition*, 19, 785-798.
- Marteniuk, R. G., Leavitt, J. L., MacKenzie, C. L., & Athenes, S. (1990). Functional relationships between grasp and transport components in a prehension task. *Human Movement Science*, 9, 149-176.
- Marteniuk, R. G., MacKenzie, C. L., Jeannerod, M., Athenes, S., & Dugas, C. (1987). Constraints on human arm movement trajectories. *Canadian Journal of Psychology*, 41, 365-378.
- Mason, A. H. (2007). Performance of Unimanual and Bimanual Multiphased Prehensile Movements. *Journal of Motor Behavior*, 39, 291-305.
- McIntyre, J., Stratta, F., & Lacquaniti, F. (1997). Viewer-Centered Frame of Reference for Pointing to Memorized Targets in Three-Dimensional Space. *Journal of Neurophysiology*, 78, 1601-1618.
- Meary, D., Chary, C., Palluel-Germain, R., & Orliaguet, J.-P. (2005). Visual perception of writing and pointing movements. *PERCEPTION-LONDON-*, 34, 1061.
- Messier, J., & Kalaska, J. F. (1999). Comparison of variability of initial kinematics and endpoints of reaching movements. *Experimental Brain Research*, 125, 139–152.
- Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, 14, 2135–2137.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, 9, 1265–1279.
- Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, 31, 236-250.
- Milner, T. E., & Ijaz, M. M. (1990). The effect of accuracy constraints on three-dimensional

- movement kinematics. *Neuroscience*, 35, 365-374.
- Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience & Biobehavioral Reviews*, 36, 341-349.
- Mori, M. (1970). The uncanny valley. *Energy*, 7, 33-35.
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-Neuron Responses in Humans during Execution and Observation of Actions. *Current Biology*, 20, 750-756.
- Mulder, T. (2007). Motor imagery and action observation: cognitive tools for rehabilitation. *Journal of neural transmission*, 114, 1265-1278.
- Müsseler, J. (1999). How independent from action control is perception?: An event-coding account for more equally-ranked crosstalks. *Advances in psychology*, 129, 121-147.
- Nagai, Y., Kawai, Y., & Asada, M. (2011). Emergence of mirror neuron system: Immature vision leads to self-other correspondence. In *Development and Learning (ICDL), 2011 IEEE International Conference on* (Vol. 2, p. 1-6).
- Naish, K. R., Reader, A. T., Houston-Price, C., Bremner, A. J., & Holmes, N. P. (2013). To eat or not to eat? Kinematics and muscle activity of reach-to-grasp movements are influenced by the action goal, but observers do not detect these differences. *Experimental Brain Research*, 225, 261-275.
- Naito, E., Kochiyama, T., Kitada, R., Nakamura, S., Matsumura, M., Yonekura, Y., & Sadato, N. (2002). Internally simulated movement sensations during motor imagery activate cortical motor areas and the cerebellum. *The Journal of neuroscience*, 22, 3683-3691.
- Neggers, S. F. W., & Bekkering, H. (2000). Ocular gaze is anchored to the target of an ongoing pointing movement. *Journal of Neurophysiology*, 83, 639-651.
- Nissen, S. (2003). Implementation of a fast artificial neural network library (fann). *Report, Department of Computer Science University of Copenhagen (DIKU)*, 31.
- Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron

- system: A link between action observation and social skills. *Social Cognitive and Affective Neuroscience*, 2, 62 -66.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Orliaguet, J. P., Vaillon, S., Coello, Y., & Kandel, S. (1996). Perceptual anticipation in sequential grasping movements. In *INTERNATIONAL JOURNAL OF PSYCHOLOGY* (Vol. 31, p. 48478–48478).
- Orliaguet, J.-P., Kandel, S., & Boe, L.-J. (1997). Visual perception of motor anticipation in cursive handwriting: Influence of spatial and movement information on the prediction of forthcoming letters. *PERCEPTION-LONDON-*, 26, 905–912.
- Oztop, E., Franklin, D. W., Chaminade, T., & Cheng, G. (2005). Human–humanoid interaction: Is a humanoid robot perceived as a human? *International Journal of Humanoid Robotics*, 2, 537–559.
- Oztop, E., Wolpert, D., & Kawato, M. (2005). Mental state inference using visual control parameters. *Cognitive Brain Research*, 22, 129–151.
- Pacherie, E. (2011). Framing Joint Action. *Review of Philosophy and Psychology*, 2, 173-192.
- Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 709.
- Parsons, L. M. (2001). Integrating cognitive psychology, neurology and neuroimaging. *Acta Psychologica*, 107, 155-181.
- Pascual-Leone, A. (2001). The brain that plays music and is changed by it. *Annals of the New York Academy of Sciences*, 930, 315–329.
- Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91. doi:10.1007/BF00230027
- Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand

- movements in a natural task. *Experimental Brain Research*, 139, 266-277.
- Pierno, A. C., & Ansuini, C. (2007). Motor Intentions versus Social Intentions: One System or Multiple Systems? *PSYCHE*, 13, 2.
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, 82, B51–B61.
- Popovic, D., Stein, R. B., Namik Oguztoreli, M., Lebedowska, M., & Jonic, S. (1999). Optimal control of walking with functional electrical stimulation: a computer simulation study. *Rehabilitation Engineering, IEEE Transactions on*, 7, 69–79.
- Pozzo, T., Papaxanthis, C., Petit, J. L., Schweighofer, N., & Stucchi, N. (2006). Kinematic features of movement tunes perception and action coupling. *Behavioural brain research*, 169, 75–82.
- Prinz, W., & Hommel, B. (2002). *Common mechanisms in perception and action: Attention and Performance XIX*. Oxford University Press.
- Pryde, K. M., Roy, E. A., & Campbell, K. (1998). Prehension in children and adults: the effects of object size. *Human Movement Science*, 17, 743–752.
- Quesque, F., Lewkowicz, D., Delevoye-Turrell, Y. N., & Coello, Y. (2013). Effects of social intention on movement kinematics in cooperative actions. *Frontiers in Neurorobotics*, 7, 14.
- Ramachandran, V. S. (2000). Mirror neurons and imitation learning as the driving force behind « the great leap forward » in human evolution. *Edge Website article* [http://www.edge.org/3rd\\_culture/ramachandran/ramachandran\\_p1.html](http://www.edge.org/3rd_culture/ramachandran/ramachandran_p1.html).
- Rand, M. K., Alberts, J., Stelmach, G., & Bloedel, J. (1997). The influence of movement segment difficulty on movements with two-stroke sequence. *Experimental brain research*, 115, 137–146.
- Rand, M. K., & Stelmach, G. E. (2000). Segment interdependency and difficulty in two-stroke sequences. *Experimental brain research*, 134, 228–236.
- Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., ... Pietrini, P.

- (2009). Do We Really Need Vision? How Blind People « See » the Actions of Others. *The Journal of Neuroscience*, 29, 9719-9724.
- Rizzolatti, G., & Fabbri-Destro, M. (2009). Mirror neurons: from discovery to autism. *Experimental Brain Research*, 200, 223-237.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131-141.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, 111. doi:10.1007/BF00227301
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci*, 2, 661-670.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2009). The mirror neuron system: A motor-based mechanism for action and intention understanding. *The Cognitive Neuroscience IV*, 625–640.
- Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, 31, 889-901.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci*, 11, 264-274.
- Rosenbaum, D. A., Hindorff, V., & Munro, E. M. (1987). Scheduling and programming of rapid finger sequences: Tests and elaborations of the hierarchical editor model. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 193.
- Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., & van der Wel, R. (2012). Cognition, action, and object manipulation. *Psychological Bulletin*, 138, 924-946.
- Rosenbaum, D. A., Slotta, J. D., Vaughan, J., & Plamondon, R. (1991). Optimal movement selection. *Psychological Science*, 2, 86–91.
- Rosenbaum, David A., & Jorgensen, M. J. (1992). Planning macroscopic aspects of manual control. *Human Movement Science*, 11, 61-69.
- Rosenbaum, David A., Marchak, F., Barnes, H. J., Vaughan, J., Slotta, J. D., & Jorgensen, M.

- J. (1990). Constraints for action selection: Overhand versus underhand grips. In *Attention and performance 13: Motor representation and control* (p. 321-342). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Rossetti, Y., Stelmach, G., Desmurget, M., Prablanc, C., & Jeannerod, M. (1994). The effect of viewing the static hand prior to movement onset on pointing kinematics and variability. *Experimental Brain Research*, 101, 323-330.
- Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of neurophysiology*, 96, 1358–1369.
- Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G. G., Matelli, M., & Luppino, G. (2006). Cortical Connections of the Inferior Parietal Cortical Convexity of the Macaque Monkey. *Cerebral Cortex*, 16, 1389 -1417.
- Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G., & Fogassi, L. (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *European Journal of Neuroscience*, 28, 1569-1588.
- Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental brain research*, 152, 229–242.
- Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Consciousness and Cognition*, 18, 766-772.
- Sartori, L., Becchio, C., Bulgheroni, M., & Castiello, U. (2009). Modulation of the action control system by social intention: unexpected social requests override preplanned action. *Journal of Experimental Psychology. Human Perception and Performance*, 35, 1490-1500.
- Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, 119, 242-252.



- Saunier, G., Papaxanthis, C., Vargas, C. D., & Pozzo, T. (2008). Inference of complex human motion requires internal models of action: behavioral evidence. *Experimental Brain Research*, 185, 399–409.
- Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2012). The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social cognitive and affective neuroscience*, 7, 413–422.
- Scassellati, B. (2002). Theory of mind for a humanoid robot. *Autonomous Robots*, 12, 13–24.
- Schneegans, S., & Schöner, G. (2008). Dynamic field theory as a framework for understanding embodied cognition. *Handbook of cognitive science: An embodied approach*, 241–271.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4, 299-309.
- Schöner, G. (2008). *Dynamical systems approaches to cognition*. Cambridge, UK: Cambridge University Press.
- Schubö, A., Prinz, W., & Aschersleben, G. (2004). Perceiving while acting: Action affects perception. *Psychological Research*, 68, 208–215.
- Schuboe, A., Maldonado, A., Stork, S., & Beetz, M. (2008). Subsequent actions influence motor control parameters of a current grasping action. In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on* (p. 389–394).
- Searle, J. R. (1983). *Intentionality, an essay in the philosophy of mind*. Cambridge University.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*, 10, 70–76.
- Seegelke, C., Hughes, C. M. L., Schütz, C., & Schack, T. (2012). Individual differences in motor planning during a multi-segment object manipulation task. *Experimental Brain Research*, 222, 125-136.
- Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during

- learning of a motor task. *The Journal of Neuroscience*, 14, 3208–3224.
- Shadmehr, R., & Wise, S. P. (2005). *The computational neurobiology of reaching and pointing: a foundation for motor learning*. MIT press.
- Short, M. W., & Cauraugh, J. H. (1997). Planning macroscopic aspects of manual control: End-state comfort and point-of-change effects. *Acta Psychologica*, 96, 133–147.
- Short, M. W., & Cauraugh, J. H. (1999). Precision hypothesis and the end-state comfort effect. *Acta Psychologica*, 100, 243-252.
- Short, M. W., Fischman, M. G., & Wang, Y. T. (1996). Cinematographical analysis of movement pathway constraints in rapid target-striking tasks. *Journal of motor behavior*, 28, 157–163.
- Sidaway, B., Sekiya, H., & Fairweather, M. (1995). Movement Variability as a Function of Accuracy Demand in Programmed Serial Aiming Responses. *Journal of Motor Behavior*, 27, 67-76.
- Sirigu, A., Cohen, L., Duhamel, J. R., & Pillon, B. (1995). Congruent unilateral impairments for real and imagined hand movements. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience; Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*.
- Sisbot, E. A., Marin-Urias, L. F., Broquere, X., Sidobre, D., & Alami, R. (2010). Synthesizing robot motions adapted to human presence. *International Journal of Social Robotics*, 2, 329–343.
- Soechting, J. F., & Flanders, M. (1992). Organization of sequential typing movements. *Journal of neurophysiology*, 67, 1275–1290.
- Soechting, J. F., Lacquaniti, F., & Terzuolo, C. A. (1986). Coordination of arm movements in three-dimensional space. Sensorimotor mapping during drawing movement. *Neuroscience*, 17, 295-311.
- Sosnik, R., Flash, T., Hauptmann, B., & Karni, A. (2007). The acquisition and implementation of the smoothness maximization motion strategy is dependent on

- spatial accuracy demands. *Experimental brain research*, 176, 311–331.
- Sosnik, R., Hauptmann, B., Karni, A., & Flash, T. (2004). When practice leads to co-articulation: the evolution of geometrically defined movement primitives. *Experimental Brain Research*, 156, 422–438.
- Stadler, W., Springer, A., Parkinson, J., & Prinz, W. (2012). Movement kinematics affect action prediction: comparing human to non-human point-light actions. *Psychological research*, 76, 395–406.
- Stapel, J. C., Hunnius, S., & Bekkering, H. (2012). Online prediction of others' actions: the contribution of the target object, action context and movement kinematics. *Psychological research*, 1–12.
- Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., & Frackowiak, R. S. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology*, 73, 373–386.
- Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1625.
- Strick, P. L., Dum, R. P., & Fiez, J. A. (2009). Cerebellum and nonmotor function. *Annual review of neuroscience*, 32, 413–434.
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, 13, 283 – 286.
- Tapus, A., Mataric, M. J., & Scassellati, B. (2007). Socially assistive robotics. *IEEE Robotics and Automation Magazine*, 14, 35.
- Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7, 907-915.
- Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5, 1226–1235.
- Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta*

*psychologica*, 116, 185–203.

- Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, 110, 124-129.
- Uithol, S., Haselager, W. F., Bekkering, H., Love, B. C., McRae, K., & Sloutsky, V. M. (2008). When do we stop calling them mirror neurons. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (Vol. 1788).
- Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., ... Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences*, 105, 2209 -2213.
- Uno, Y., Kawato, M., & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological cybernetics*, 61, 89–101.
- Urgesi, C., Candidi, M., Fabbro, F., Romani, M., & Aglioti, S. M. (2006). Motor facilitation during action observation: topographic mapping of the target muscle and influence of the onlooker's posture. *European Journal of Neuroscience*, 23, 2522–2530.
- Van Beers, R. J., Haggard, P., & Wolpert, D. M. (2004). The role of execution noise in movement variability. *Journal of Neurophysiology*, 91, 1050–1063.
- Van Beers, R. J., Sittig, A. C., & Denier van der Gon, J. J. (1998). The precision of proprioceptive position sense. *Experimental Brain Research*, 122, 367–377.
- Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, 48, 564–584.
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23, 998-1003.
- Vesper, C., van der Wel, R., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211, 517-530.
- Vindras, P., & Viviani, P. (1998). Frames of reference and control parameters in visuomanual pointing. *Journal of experimental psychology. Human perception and performance*,

- Vindras, P., & Viviani, P. (2005). Planning short pointing sequences. *Experimental brain research, 160*, 141–153.
- Viviani, P., & Schneider, R. (1991). A developmental study of the relationship between geometry and kinematics in drawing movements. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 198.
- Viviani, P., & Terzuolo, C. A. (1983). The organization of movement in handwriting and typing. *Language production, 2*, 103–146.
- Viviani, Paolo. (2002). Motor competence in the perception of dynamic events. *Common mechanisms in perception and action: attention and performance, 19*, 406–442.
- Webb, A., Knott, A., & MacAskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta psychologica, 133*, 51–56.
- Weiss, P. H., Jeannerod, M., Paulignan, Y., & Freund, H.-J. (2000). Is the organisation of goal-directed action modality specific? A common temporal structure. *Neuropsychologia, 38*, 1136-1147.
- Weiss, P., Stelmach, G. E., & Hefter, H. (1997). Programming of a movement sequence in Parkinson's disease. *Brain: A Journal of Neurology, 120 ( Pt 1)*, 91-102.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological bulletin, 131*, 460.
- Wing, A. M., Turton, A., & Fraser, C. (1986). Grasp size and accuracy of approach in reaching. *Journal of Motor Behavior, 18*, 245-260.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*, 593–602.
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *nature neuroscience, 3*, 1212–1217.
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor

- control. *Neural Networks*, 11, 1317–1329.
- Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2, 338-347.
- Won, J., & Hogan, N. (1995). Stability properties of human reaching movements. *Experimental Brain Research*, 107, 125-136.
- Zaal, F. T. J. M., & Bootsma, R. J. (1993). Accuracy demands in natural prehension. *Human Movement Science*, 12, 339-345.
- Zajonc, R. B. (1965). *Social facilitation*. Research Center for Group Dynamics, Institute for Social Research, University of Michigan.
- Zentgraf, K., Munzert, J., Bischoff, M., & Newman-Norlund, R. D. (2011). Simulation during observation of human actions – Theories, empirical studies, applications. *Vision Research*, 51, 827-835.
- Zijdewind, I., Toering, S. T., Bessem, B., van der Laan, O., & Diercks, R. L. (2003). Effects of imagery motor training on torque production of ankle plantar flexor muscles. *Muscle & nerve*, 28, 168–173.

# Abstract

The aim of my PhD thesis was to participate in the construction of a new humanoid robot that can sustain intuitive interactions with humans through observation and imitation. As such, I conducted a series of experimental studies in young adults to better characterize the kinematic properties of those arm movements performed during motor and social interactions, elements that are the reference patterns for the to-come robot.

Focusing on non-verbal behavior, we tested how external and internal constraints (difficulty, predictability, timing) shaped the kinematics of both arm and hand movements in a very simple pick and place sequential action (study 1 and 2). The results revealed early modulations in kinematics in the reach-to-grasp phase depending of the size and the stability of the target pad on which the object had to be placed. These modulations observed within the first element of the sequence were in contradiction with the current optimized trajectory models used in robotics for action sequences. They suggest in fact a strong coupling of the motor parameters within an encapsulated planning strategy that back-propagates the contextual constraints on to the early elements of the motor sequence. To confirm these findings, a second series of studies were conducted using kinematic and video based tasks to show that human motor intentions can be read through the detection of these early kinematic modulations (study 3). Using basic artificial classification, we tested whether low-level motor indices could afford trial categorization without the need for higher-level process such as motor imagery. Results indicated that indeed without cognitive abilities the neural network could categorize the intention of an observed action within the first 500ms, significantly above chance level (study 4). The third phase of my PhD work turned to eye tracking. Here, we revealed that the proactive strategy of eye-fixations used during action observation were similar to those made during executed actions. Additionally, good categorization of motor intention was characterized by more accurate saccades and longer object fixations. Eye movements are known to play an important role in social interactions. Hence, in a final experiment (study 5), we setup a face-to-face competitive game to reveal the specific effects that the social context may play on the kinematic properties of reaching during different types of interactive situations.

In the PhD manuscript, I propose a general discussion that sets these results within the current cognitive and neuronal models of sensori-motor integration. When appropriate, future directions are suggested both for cognitive models of motor control and for the development of neuro-inspired artificial systems constituted with intuitive social interaction skills.

Keywords : Motor Planning, Social Interaction, Action and Goals Understanding, Sequential Actions, Kinematics, Eye-Tracking, Computational models.